



SISTEMA DE CLASIFICACIÓN AUTOMÁTICA DE IMÁGENES MÉDICAS BASADO EN CONTENIDO

**Proyecto Fin de Carrera
Ingeniería de Telecomunicación**

UNIVERSIDAD CARLOS III DE MADRID

ESCUELA POLITÉCNICA SUPERIOR

Autor: Natalia Muñoz Sarriá
Tutor: Julio Villena Román
Madrid, 18 de Octubre 2010

Título: Sistema de clasificación de imágenes médicas basado en contenido

Autor: Natalia Muñoz Sarriá

Tutor: Julio Villena Román

EL TRIBUNAL

Presidente:

Secretario:

Vocal:

Realizado el acto de defensa y lectura del Proyecto de Fin de Carrera el día en Leganés, en la Escuela Politécnica Superior de la Universidad Carlos III de Madrid, acuerda otorgarle la calificación de:

Fdo. Presidente

Fdo. Secretario

Fdo. Vocal

Resumen

El objetivo del proyecto es el análisis, diseño e implementación de un sistema de anotación automática de imágenes médicas mediante un código estandarizado, el código IRMA, que aporta información técnica e información anatómica acerca de la parte del cuerpo humano analizada y del sistema biológico al que se refiere la imagen. Al igual que todos los sistemas de análisis de imágenes por contenido (CBIR), este sistema de clasificación de imágenes se basa en dos módulos, uno de extracción de características y otro de clasificación. El módulo de extracción de características obtiene la información de cada imagen y el módulo de clasificación permite, a partir de las características de la imagen obtenidas en el módulo anterior, colocar la imagen dentro de una clase, proporcionándole una codificación.

Para esta implementación se realiza un estudio de diferentes características que se pueden obtener de las imágenes y de un algoritmo de aprendizaje que permita clasificarlas después de obtener su contenido.

El desarrollo de este tipo de sistemas tiene dos fases, una de entrenamiento con imágenes previamente clasificadas que se usan para dar “inteligencia” al sistema, y otra de test con imágenes cuya clasificación se desconoce para evaluar la eficacia de clasificación del mismo.

Palabras clave

Agradecimientos

En primer lugar, agradecerte la atención a ti por haber abierto este proyecto, que por unos motivos u otros ha llamado tu atención, espero que te ayude a encontrar lo que estás buscando.

Ahora, de forma no tan abstracta, me gustaría dar las gracias a mis santos (padres), Eugenio y Maribel, por su infinita paciencia (a veces no tan infinita), por su apoyo incondicional, por tener siempre las palabras adecuadas (aunque no siempre sean las que se quieren oír) y sobre todo por todos los consejos, que aunque quizá no he sabido interpretarlos al principio, han cogido forma con el paso del tiempo.

A Fer, por estar desde el principio, y por estar siempre, que no es lo mismo. Por todas las veces que me ha llevado de la mano (tanto en sentido literal como figurado) y por las muchas que me ha ayudado a levantarme, por hacerme reír y por hacerme llorar, por recordarme que la vida es en color, y no en blanco y negro, por las veces que me ha escuchado y por las veces que me hace perder la paciencia. Por ser mis ojos cuando no veo y mis oídos cuando no quiero escuchar. Hold my hand, don't let me fall.

A mi familia, que aunque dicen que es impuesta, yo me siento muy afortunada porque en caso de haberla elegido no lo habría hecho mejor. A Lola por sus croquetas y sobre todo porque aunque no soy la mejor nieta del mundo (sin duda) ella sí es la mejor abuela; a Luis y Macovi porque más que mi familia, son mis amigos; a Kurro, a Sara y a Iker, por dar el punto de color a cada reunión.

A Rosa, por aguantarme día a día, por sus palabros (que ya son de todos y pronto estarán en la RAE), por su risa que te llega a absorber el alma, por su estrés que hace que sincronices tu reloj, por sus tacones en el silencio de la biblioteca. Por ser una amiga en todo momento, gracias.

A Inés, porque no tengo recuerdos en los que no aparezca, y por eso mismo no me imagino una vida en la que no esté, por los jueves y el vino, por su oh de la vida y por partirse y mondar siempre, por hacerme compañía sin hablar y por hacerme sentir parte de una gran familia. Brindo por ello.

A Eli, Jorge, Mamen y Paco, por ser tan distintos e intentar siempre buscar cosas que hacer juntos (aunque no siempre es fácil), y por haber entendido que la amistad no entiende ni de tiempo ni de distancia, por aprender día a día a aceptarnos y querernos, por ser incondicionales, por el parque y las pipas que se convirtieron en terraza y barbacoa (pasando por tapas y chino). Por saber envejecer tan bien (unos mejor que otros).

A Vaillo, Miki y Dani, por ser los hermanos que no tengo y los magos de la piedra filosofal. Por los días y las noches (y por las tardes, por qué no...) y por las caminatas a ninguna parte (gracias Vaillo). Por La Escapada, las plazas de Madrid, los "viajecitos" en verano (y en invierno), y sobre todo por ser tan guasones y saber poner el hombro cuando es necesario. Nunca quemaré vuestros recuerdos, porque sois unos golfos que no paran de salir, beber, el rollo de siempre, y sobre todo por amar, amar, amar y ensanchar mi alma.

A Javo (se siempre feliz) y a Manu (Jumanji?), entre otras muchas cosas, por ser parte de los Von Trapp.

A Bea y Raul, a los nuevos y a los antiguos (ambos tienen su encanto), por hacerme creer cuando estoy apunto de perder la fé. Para mí, sois un ejemplo.

A Rober, Santi e Irene, por los muchos momentos y las miles de conversaciones, que no se acaben nunca porque aún quedan muchas cosas que contar (eso que no falte).

A Kiko, Recu y Carlos, por completar un grupo (espero no dejarme a nadie).

A mi gente de Düsseldorf, amigos y compañeros, familia fuera de la madre patria.

A Julio, porque aunque no he tenido más tutores de proyecto me llevo la grata sensación de que es, cuanto menos, diferente. Muchas gracias por la paciencia inicial (empezar es... difícil...), por el enorme esfuerzo de los últimos días, por los consejos y por su apoyo.

ÍNDICE

1.- INTRODUCCIÓN.....	1
1.1.- MOTIVACIÓN DEL PROYECTO	1
1.2.- OBJETIVOS DEL PROYECTO	2
1.3.- ESTRUCTURA DEL DOCUMENTO.....	3
2.- ESTADO DEL ARTE	5
2.1.- INTELIGENCIA ARTIFICIAL	5
2.1.1.- <i>Introducción a la Inteligencia Artificial</i>	5
2.1.2.- <i>Inteligencia Artificial en el diagnóstico médico</i>	6
2.1.3.- <i>Sistemas expertos</i>	9
2.1.4.- <i>Lógica difusa o borrosa</i>	11
2.1.5.- <i>Redes neuronales</i>	12
2.1.6.- <i>Minería de datos aplicada</i>	13
2.2.- SISTEMAS DE CLASIFICACIÓN AUTOMÁTICA DE IMÁGENES BASADOS EN CONTENIDO	14
2.3.1.- <i>Introducción a los sistemas de clasificación de imágenes basados en contenido</i>	14
2.3.2.- <i>Características de las imágenes</i>	16
2.3.3.- <i>Extracción de características</i>	18
2.3.4.- <i>Clasificación y codificación de imágenes</i>	42
2.3.5.- <i>Base de datos</i>	45
2.4.- IMAGECLEF: ANOTACIÓN AUTOMÁTICA DE IMÁGENES MÉDICAS	48
3.- ANÁLISIS DEL SISTEMA DE CLASIFICACIÓN DE IMÁGENES.....	53
3.1.- INTRODUCCIÓN.....	53
3.2.- EXTRACCIÓN DE CARACTERÍSTICAS	54
3.3.- BÚSQUEDA DE IMÁGENES (ALGORITMO K-NN).....	70
4.- IMPLEMENTACIÓN DEL SISTEMA DE CLASIFICACIÓN DE IMÁGENES.....	74
4.1.- INTRODUCCIÓN.....	74
4.2.- FUNCIONES IMPLEMENTADAS.....	74
4.2.1.- <i>Funciones de procesamiento de imágenes</i>	74
4.2.2.- <i>Funciones de extracción de características</i>	77
4.2.3.- <i>Funciones de clasificación</i>	86
4.2.4.- <i>Funciones de evaluación de los resultados</i>	88
5.- DESARROLLO Y PRUEBAS	90
5.1.- INTRODUCCIÓN.....	90
5.2.- PRUEBAS REALIZADAS	91
5.2.1.- <i>Características por separado</i>	91
5.2.2.- <i>Inclusión de todas las características</i>	101
5.2.3.- <i>Eliminación de características una por una</i>	102
5.3.- CONCLUSIONES ANTE LAS PRUEBAS REALIZADAS	111
6.- EVALUACIÓN DE RESULTADOS.....	114
6.1.- INTRODUCCIÓN.....	114
6.2.- TASA DE ERROR.....	114
6.2.1.- <i>Características por separado</i>	114
6.2.2.- <i>Inclusión de todas las características</i>	115
6.2.3.- <i>Eliminación de características una por una</i>	116
6.2.4.- <i>Asignación de diferentes pesos</i>	117
6.3.- ANÁLISIS DE LOS RESULTADOS OBTENIDOS	118
6.3.1.- <i>Características</i>	118
6.3.2.- <i>Valores de k</i>	119
6.3.3.- <i>Diseño óptimo</i>	119
6.3.4.- <i>Posibles causas de error</i>	119
6.3.5.- <i>Comparación con otros sistemas (ImageCLEF 2007)</i>	123
7.- PRESUPUESTO.....	125

8.- CONCLUSIONES Y TRABAJOS FUTUROS.....	126
8.1.- INTRODUCCIÓN.....	126
8.2.- CONCLUSIONES OBTENIDAS DE LA REALIZACIÓN DEL PROYECTO.....	126
8.3.- TRABAJOS FUTUROS	128
APÉNDICES.....	132
APÉNDICE A. CLASES IRMA.....	132
APÉNDICE B. CÓDIGOS IRMA.....	133
BIBLIOGRAFÍA Y REFERENCIAS.....	141

ÍNDICE DE FIGURAS

FIGURA 1. EJEMPLO HISTOGRAMA.....	19
FIGURA 2. OPERACIÓN MORFOLÓGICA DE DILATACIÓN.....	31
FIGURA 3. OPERACIÓN MORFOLÓGICA DE EROSIÓN.....	32
FIGURA 4. FILTRO DE APERTURA.....	33
FIGURA 5. FILTRO DE CIERRE.....	33
FIGURA 6. NIVELES RELEVANTES DE LOS COEFICIENTES DE LA TRANSFORMADA DISCRETA DEL COSENO ...	38
FIGURA 7. MATRIZ OBTENIDA DE LA TRANSFORMADA DISCRETA DEL COSENO.....	39
FIGURA 8. REORDENACIÓN EN ZIG-ZAG DE LA TRANSFORMADA DISCRETA DEL COSENO.....	39
FIGURA 9. DIAGRAMA DE BLOQUES DE LA TRANSFORMADA DISCRETA DE WAVELET.....	41
FIGURA 10. ESQUEMA DE COEFICIENTES DE WAVELET.....	42
FIGURA 11. DIAGRAMA DE BLOQUES DEL PROCESO DE CLASIFICACIÓN Y CODIFICACIÓN DE IMÁGENES.....	43
FIGURA 12. MATRIZ DE CARACTERÍSTICAS.....	44
FIGURA 13. IMÁGENES DE EJEMPLO.....	54
FIGURA 14. IMAGEN DE EJEMPLO ALGORITMO K-NN.....	71
FIGURA 15. IMAGEN EJEMPLO PARA PRUEBAS.....	91
FIGURA 16. VARIABILIDAD EN IMÁGENES DEL CÓDIGO 1121-115-700-400.....	111
FIGURA 17. VARIABILIDAD EN IMÁGENES DEL CÓDIGO 1121-120-463-700.....	111
FIGURA 18. ERROR EN EL EJE DE DIRECCIÓN.....	112
FIGURA 19. ERROR EN EL EJE ANATÓMICO.....	112
FIGURA 20. ERROR EN IMÁGENES GIRADAS.....	113
FIGURA 21. REPRESENTACIÓN DE LA TASA DE ERROR DE CADA CARACTERÍSTICA PARA DISTINTOS VALORES DE K.....	115
FIGURA 22. REPRESENTACIÓN DE LA TASA DE ERROR PARA TODAS LAS CARACTERÍSTICAS Y DISTINTOS VALORES DE K.....	116
FIGURA 23. REPRESENTACIÓN DE LA TASA DE ERROR ELIMINANDO UNA CARACTERÍSTICA Y DISTINTOS VALORES DE K.....	117
FIGURA 24. NÚMERO DE INDIVIDUOS DE CADA CLASE.....	120
FIGURA 25. TASA DE ERROR (%) DE LAS CLASES 1 A 58.....	122
FIGURA 26. TASA DE ERROR (%) DE LAS CLASES 59 A 116.....	123
FIGURA 27. COMPARACIÓN DEL SISTEMA CON IMAGECLEF 2007.....	123

ÍNDICE DE TABLAS

TABLA 1. MÁSCARAS DE FILTROS BIDIRECCIONALES DE LAWS.....	24
TABLA 2. IMÁGENES DESPUÉS DEL ENVENTANADO	24
TABLA 3. HISTOGRAMA DE IMÁGENES DE EJEMPLO	55
TABLA 4. ESTADÍSTICOS DE IMÁGENES DE EJEMPLO	56
TABLA 5. DIMENSIÓN FRACTAL DE IMÁGENES DE EJEMPLO.....	57
TABLA 6. FILTROS DE GABOR PARA 24 COMBINACIONES DE ESCALAS Y ORIENTACIONES.....	58
TABLA 7. CARACTERÍSTICAS DE GABOR DE IMÁGENES DE EJEMPLO	60
TABLA 8. IMÁGENES FILTRADAS CON MÁSCARAS DE LAWS DEL EJEMPLO 1	61
TABLA 9. IMÁGENES FILTRADAS CON MÁSCARAS DE LAWS DEL EJEMPLO 2.....	62
TABLA 10. IMÁGENES FILTRADAS CON MÁSCARAS DE LAWS DE EJEMPLO 3.....	62
TABLA 11. IMÁGENES FILTRADAS CON FILTROS DE LAWS DEL EJEMPLO 4	63
TABLA 12. IMÁGENES FILTRADAS CON FILTROS LAWS DEL EJEMPLO 5	63
TABLA 13. ENERGÍA DE LAS IMÁGENES FILTRADAS CON MATRICES DE CARACTERÍSTICAS SIMILARES.....	64
TABLA 14. CARACTERÍSTICAS DE TAMURA DE LAS IMÁGENES DE EJEMPLO.....	65
TABLA 15. ESTADÍSTICOS DE LA MATRIZ DE CO-OCURRENCIA DE LAS IMÁGENES DE EJEMPLO	66
TABLA 16. COEFICIENTES DE LA TRANSFORMADA DISCRETA DEL COSENO DE LAS IMÁGENES DE EJEMPLO.....	68
TABLA 17. ESTADÍSTICOS DE LA TRANSFORMADA DISCRETA DE WAVELET DE LAS IMÁGENES DE EJEMPLO.....	68
TABLA 18. MÓDULO DE LOS COEFICIENTES DE LA TRANSFORMADA DISCRETA DE FOURIER DE LAS IMÁGENES DE EJEMPLO	69
TABLA 19. CARACTERÍSTICAS DE CADA IMAGEN	70
TABLA 20. RESULTADOS ALGORITMO K-NN DE IMAGEN DE EJEMPLO.....	73
TABLA 21. RESULTADOS DE CLASIFICACIÓN PARA EL HISTOGRAMA	92
TABLA 22. RESULTADOS DE CLASIFICACIÓN PARA LOS ESTADÍSTICOS	93
TABLA 23. RESULTADOS DE CLASIFICACIÓN PARA LA DIMENSIÓN FRACTAL.....	94
TABLA 24. RESULTADOS DE CLASIFICACIÓN PARA LAS CARACTERÍSTICAS DE GABOR.....	95
TABLA 25. RESULTADOS DE CLASIFICACIÓN PARA LAS CARACTERÍSTICAS DE TAMURA	96
TABLA 26. RESULTADOS DE CLASIFICACIÓN PARA LA MATRIZ DE CO-OCURRENCIA.....	97
TABLA 27. RESULTADOS DE CLASIFICACIÓN PARA LA TRANSFORMADA DISCRETA DEL COSENO, DCT	98
TABLA 28. RESULTADOS DE CLASIFICACIÓN PARA LA TRANSFORMADA DISCRETA DE WAVELET, DWT.....	99
TABLA 29. RESULTADOS DE CLASIFICACIÓN PARA LA TRANSFORMADA DISCRETA DE FOURIER, DFT	100
TABLA 30. RESULTADOS DE CLASIFICACIÓN PARA TODAS LAS CARACTERÍSTICAS.....	101
TABLA 31. RESULTADOS DE CLASIFICACIÓN ELIMINANDO EL HISTOGRAMA	102
TABLA 32. RESULTADOS DE CLASIFICACIÓN ELIMINANDO LOS ESTADÍSTICOS	103
TABLA 33. RESULTADOS DE CLASIFICACIÓN ELIMINANDO LA DIMENSIÓN FRACTAL	104
TABLA 34. RESULTADOS DE CLASIFICACIÓN ELIMINANDO LAS CARACTERÍSTICAS DE GABOR	105
TABLA 35. RESULTADOS DE CLASIFICACIÓN ELIMINANDO LAS CARACTERÍSTICAS DE TAMURA	106
TABLA 36. RESULTADOS DE CLASIFICACIÓN ELIMINANDO LA MATRIZ DE CO-OCURRENCIA.....	107
TABLA 37. RESULTADOS DE CLASIFICACIÓN ELIMINANDO LA TRANSFORMADA DISCRETA DEL COSENO, DCT	108
TABLA 38. RESULTADOS DE CLASIFICACIÓN ELIMINANDO LA TRANSFORMADA DISCRETA DE WAVELET, DWT	109
TABLA 39. RESULTADOS DE CLASIFICACIÓN ELIMINANDO LA TRANSFORMADA DISCRETA DE FOURIER, DFT	110
TABLA 40. TASA DE ERROR DE CADA CARACTERÍSTICA PARA DISTINTOS VALORES DE K	114
TABLA 41. TASA DE ERROR OBTENIDA CON TODAS LAS CARACTERÍSTICAS Y DIFERENTES VALORES DE K	115
TABLA 42. TASA DE ERROR OBTENIDA PARA LA ELIMINACIÓN DE UNA CARACTERÍSTICA Y DISTINTOS VALORES DE K	116
TABLA 43. PESOS DE LAS CARACTERÍSTICAS DEL ESTUDIO	118
TABLA 44. TASA DE ERROR OBTENIDA PARA EL ESTUDIO 1 Y DIFERENTES VALORES DE K	118
TABLA 45. ACIERTOS Y FALLOS PARA LAS DIFERENTES CLASES.....	122
TABLA 46. MEJORES/PEORES RESULTADOS SEGÚN INDIVIDUOS ENTRENAMIENTO	122
TABLA 47. EJEMPLO DE K CÓDIGOS IRMA DE LOS K VECINOS MÁS CERCANOS A UNA IMAGEN DE TEST ...	130
TABLA 48. EJEMPLO DE K CÓDIGOS IRMA CON LAS PROBABILIDADES POR EJE	130
TABLA 49. EJEMPLO DE N POSIBLES CÓDIGOS AL COMBINAR LOS K CÓDIGOS DE LOS VECINOS MÁS CERCANOS	130
TABLA 50. EJEMPLO DE LOS POSIBLES CÓDIGOS VALIDOS DE UNA IMAGEN DE TEST	131

1.- INTRODUCCIÓN

1.1.- Motivación del proyecto

El procesado de imágenes digitalizadas presenta una serie de ventajas frente al de las imágenes almacenadas en papel, de manera que su uso ha crecido exponencialmente en los últimos años, ya sea en la vida cotidiana con el uso de las cámaras digitales, como en otras áreas profesionales como la medicina o el periodismo. Estas ventajas son, la mejora de la información pictórica para la posterior interpretación humana y la mejora en el procesado de las imágenes para su posterior almacenamiento, transmisión y representación de las mismas en máquinas autónomas de percepción [25]. Otra gran ventaja que ofrecen estas imágenes es la gran capacidad de almacenamiento de información en un espacio físico reducido. Esto ha hecho que el volumen de imágenes generado sea tan grande, que haya sido necesaria la implementación de sistemas que permitan clasificar estas imágenes de forma automática, para posteriormente almacenar en una base de datos que permita realizar consultas de manera rápida y sencilla.

Particularmente en el área médica, ya que este proyecto se centra en ella, se dispone de un gran número de tipos de imágenes que son usadas día a día para el diagnóstico y seguimiento de pacientes en centros hospitalarios (radiografías, tomografías, resonancias, ecografías...). La aparición de estos sistemas de clasificación y búsqueda de imágenes da soporte a esta tarea de diagnóstico y seguimiento.

El estudio y posterior clasificación de una imagen no puede hacerse de forma automatizada sin la colaboración de especialistas que puedan asociar cada imagen a un conjunto determinado. Este proceso de clasificación y etiquetado de imágenes es una tarea tediosa y además proporciona unos resultados imprecisos y subjetivos a la persona que lo ha realizado. Además, para los seres humanos es imposible expresar alguna de las características visuales de las imágenes en lenguaje natural.

Otro problema de los sistemas de búsqueda basados en texto, a parte de la tediosidad de la tarea de etiquetado y la subjetividad de la misma, es el idioma en el que estén escritas las etiquetas, ya que el usuario puede estar buscando un mismo concepto en un idioma y la imagen puede estar etiquetada en otro distinto. Pero aún así, con las limitaciones que hemos descrito de estos sistemas, Google Image sigue siendo el buscador de imágenes más potente que existe en la actualidad.

Fue debido a estas limitaciones de los sistemas de búsqueda por texto por lo que surgieron los sistemas de clasificación de imágenes basados en contenido, que no necesitan un etiquetado previo de las imágenes para su clasificación y la posterior realización de consultas. Estos sistemas se basan en la extracción de un conjunto de características de las

imágenes (que puede variar dependiendo de la finalidad con que se haga o del tipo de imágenes que se este analizando); estas características pueden ser de tipo morfológico y/o estructural. Estos sistemas se denominan CBIR (*Content Based Image Retrieval*) y se pueden usar independientemente, o como complemento de un sistema de búsqueda de imágenes por texto, aumentando de esta manera sus prestaciones.

Una de sus mayores ventajas es la capacidad de búsqueda de una imagen sin necesidad de proporcionar al sistema una descripción de la misma, sino que se basa en características visuales que se extraen a partir de ciertos parámetros obtenidos directamente de la imagen, y otros tras haberla sometido a distintos tratamientos (transformadas, filtros, máscaras...). Estas características se comparan con las de las imágenes almacenadas previamente en la base de datos y posteriormente se aplica un decisor, que podrá basarse en la medida de distintos tipos de distancias, dependiendo de la finalidad del resultado y del tipo de imagen que se esté tratando.

Además de elegir el tipo de distancia usada en el decisor, también se puede dar un peso distinto a cada una de las características dependiendo de la importancia que tengan para la aplicación concreta en que las se estén utilizando.

Debido al incremento de estudios sobre este tipo de sistemas de clasificación de imágenes basados en contenido, se ha decidido enfocar este proyecto al diseño y posterior implementación de uno de ellos, concretamente orientado a la clasificación de imágenes médicas (aunque pueden ser utilizados en multitud de aplicaciones), usando como método de codificación de las mismas el código estándar IRMA (*Image Retrieval in Medical Applications*). Este tipo de codificación se basa en un sistema ortogonal de cuatro ejes con tres o cuatro posiciones. Cada uno de estos ejes da un tipo de información independiente de la que da el resto, de forma jerárquica. La información que ofrece esta codificación es:

Técnica: tipo de imagen y forma de adquisición.

De dirección: orientación de la imagen con respecto al cuerpo.

Anatómica: región del cuerpo examinada.

Biológica: sistema biológico examinado.

1.2.- Objetivos del proyecto

El objetivo de este proyecto es el diseño, implementación y posterior evaluación de un sistema de clasificación de imágenes basado en contenido. Para entrenar al sistema se utilizará una base de datos de 10.000 imágenes de las que se conoce el código (conjunto de entrenamiento) y para evaluarlo otra base de datos de 1.000 imágenes cuyo código se desconoce (conjunto de test), de manera que el sistema asigne el código IRMA (*Image Retrieval in Medical Applications*) a cada una de las imágenes dadas.

La iniciativa de este Proyecto de Fin de Carrera se remonta a la participación en años sucesivos, de distintos miembros del Departamento de Telemática de la Universidad Carlos III de Madrid en CLEF (*Cross Language Evaluation Forum*), un evento en el que se evalúan sistemas de recuperación de información multilingüe y multimedia desde el año 2000.

En definitiva, con este proyecto se pretende clasificar un conjunto de imágenes denominadas “de test” a partir de la extracción de sus características más importantes (tras un estudio previo para identificar qué características o conjunto de características proporciona mayor información) mediante su codificación a través de un código estándar, el código IRMA, de manera que permita posteriormente la realización de consultas y recuperación de imágenes similares a una previamente seleccionada o partir de la descripción de una serie de características de la misma.

Para la clasificación de las imágenes se ha estudiado un gran volumen de características y combinaciones de las mismas, tanto de forma global en la imagen completa, como de forma local extrayendo características de subimágenes en las que se ha dividido previamente la imagen completa, siendo estas características locales más robustas frente a traslaciones y otros cambios que las que se obtienen de forma global. Esta división de la imagen en subimágenes más pequeñas puede hacerse tanto de forma sistemática, dividiendo la imagen en un número de partes de un tamaño concreto (probando la eficacia del sistema con diferentes tamaños de estas subimágenes), como de forma más concreta buscando patrones previamente establecidos.

El diseño del sistema se ha hecho de tal manera que sea flexible y modular, para que puedan extenderse sus aplicaciones en un futuro hacia otro tipo de imágenes que no sean precisamente médicas. Además se ha buscado la sencillez en el mismo de manera que no sean precisos conocimientos médicos muy profundos para la realización de consultas.

1.3.- Estructura del documento

La presente memoria se divide en ocho capítulos:

El primer capítulo da una breve visión sobre el “por qué” de este proyecto (la motivación) y el “qué” del mismo (los objetivos), así como un resumen de la estructura del documento completo.

El segundo capítulo es la introducción teórica en la que se basa el posterior diseño e implementación del sistema de clasificación de imágenes médicas basado en contenido. Comienza por un resumen de la historia de la Inteligencia Artificial y sus investigaciones

centradas en el diagnóstico médico; posteriormente hace un breve repaso de los distintos métodos de recuperación y clasificación de imágenes (sistemas expertos, lógica difusa, redes neuronales y minería de datos aplicada); y finalmente explica en detalle los sistemas de clasificación de imágenes basados en contenido (extracción de características, clasificación, base de datos y aplicaciones).

El tercer capítulo describe los diferentes módulos en los que se divide el sistema de clasificación de imágenes diseñado, haciendo especial énfasis en el módulo de extracción de características y en el de búsqueda de imágenes. Para el estudio de las diferentes características se han seleccionado cuatro imágenes, entre las que se encuentran dos que podrían considerarse parecidas entre sí, y otras dos que no tienen nada que ver ni con las dos anteriores ni entre sí. De esta manera podemos ver que características nos permiten diferenciar las imágenes de una manera más eficaz.

El cuarto capítulo habla de forma técnica de los módulos de los que se compone el sistema y las funciones implementadas con que se ha desarrollado (funciones de extracción de características, funciones de obtención de imágenes y funciones de evolución de resultados).

En el quinto capítulo se describen las pruebas realizadas para comprobar el funcionamiento del sistema, los problemas encontrados durante el diseño e implementación del mismo y los cambios que se han ido realizando para mejorar los resultados.

En el sexto capítulo se exponen los resultados obtenidos con el diseño final del sistema y alguno de los pasos intermedios hasta llegar a él para ver qué parámetros varían estos resultados y de qué manera. Para llegar a estos resultados se han cogido una imagen de test y tres de entrenamiento y se comparan los valores de las características y de la tasa de error obtenida para cada una de ellas (medidas de similitud). Finalmente se da una visión global de los resultados una vez procesadas todas las imágenes del conjunto de imágenes de test.

En el séptimo capítulo se detalla el presupuesto del proyecto.

Finalmente, en el capítulo octavo se exponen las conclusiones del desarrollo del proyecto y los posibles trabajos futuros que podrían tener lugar a continuación.

Los dos últimos apartados contienen los apéndices, las referencias teóricas y la bibliografía en que se ha basado el estudio completo.

2.- ESTADO DEL ARTE

2.1.- Inteligencia artificial

2.1.1.- Introducción a la Inteligencia Artificial

La Inteligencia Artificial es una combinación de ciencias computacionales, fisiología y filosofía que agrupa varios campos, tales como robótica, sistemas expertos, sistemas inteligentes y muchos otros, todos los cuales tienen como objetivo común la creación de “máquinas que pueden pensar”, tal y como lo hacen los humanos [33]. La idea de construir una máquina que pueda realizar tareas que necesitan de la inteligencia humana para llevarse a cabo es un atractivo que hace de la Inteligencia Artificial uno de los principales objetivos de los científicos e investigadores de hoy en día. Las tareas que han sido estudiadas desde este punto de vista incluyen juegos, traducción de idiomas, diagnóstico (médico, de fallas,...), robótica, suministro de asesoría experta y un largo etcétera.

La investigación en las áreas de aprendizaje, lenguaje y percepción sensorial por parte de algunos científicos ha ayudado a la construcción de estas máquinas inteligentes. Uno de los puntos que más enfrenta a los expertos en el tema es la creación de sistemas que imitan el comportamiento del cerebro humano, compuesto de miles de millones de neuronas, y posiblemente el sistema más complejo que existe. La mejor manera de medir la inteligencia de una máquina es el test que propuso el informático británico Alan Turing, que dijo que un equipo merece ser llamado “*inteligente*” si es capaz de hacer creer a un ser humano que en realidad es un ser humano y no una máquina.

Los inicios de la Inteligencia Artificial se remontan a tiempos anteriores a los orígenes de la electrónica, según apuntan ciertos filósofos y matemáticos como Boole, que teorizan sobre qué principios fueron utilizados como base lógica de la Inteligencia Artificial. La aparición del ordenador, en torno a 1943, es el hecho que impulsó a los científicos a indagar en el mundo de la Inteligencia Artificial. En aquel momento se pensó que la tecnología disponible era suficiente para simular un comportamiento inteligente. Sin embargo, durante las cuatro décadas siguientes se ha demostrado que no era cierto y se ha continuado mejorando, y a pesar de muchos obstáculos, la Inteligencia Artificial ha pasado de tener una docena de investigadores, a miles de ingenieros y especialistas, dando lugar a programas capaces de jugar a las damas, e incluso sistemas diseñados para diagnosticar una enfermedad.

Hay pruebas de que los inicios de la Inteligencia Artificial se remontan al antiguo Egipto, pero fue con la aparición del ordenador, en 1943, cuando finalmente se convirtió en una tecnología disponible para crear la “máquina inteligente”, es decir, capaz de pensar, imitando el comportamiento del cerebro humano. Finalmente, el término Inteligencia Artificial se acuñó en 1956, en la conferencia de Dartmouth, y desde entonces este concepto se ha ampliado debido

a la aparición de teorías y principios desarrollados por investigadores. A través de su corta historia, el progreso en este campo ha sido más lento de lo que se preveía en un primer momento, pero se sigue avanzando día a día.

Uno de los trabajos teóricos más importantes que ha tenido lugar en las investigaciones acerca de la Inteligencia Artificial fue el desarrollo de unos algoritmos matemáticos por Warren McCullock y Walter Pitts [41], en 1943, necesarios para posibilitar el trabajo de clasificación, o funcionamiento, en sentido general, de una Red Neuronal. En 1949 Donald Hebb desarrolló un algoritmo de aprendizaje para dichas redes neuronales creando, en conjunto con los trabajos de McCullock y Pitts, la escuela creacionista. Esta escuela se considera hoy como el origen de la Inteligencia Artificial; sin embargo se trató poco por muchos años, dando paso al razonamiento simbólico basado en reglas de producción, lo que hoy en día se conoce como Sistemas Expertos.

2.1.2.- Inteligencia Artificial en el diagnóstico médico

Desde los inicios de la Inteligencia Artificial, ubicados, tal y como se ha dicho anteriormente, aproximadamente a finales de la década de los 50, sus técnicas han sido utilizadas y aplicadas en diversas áreas comunes al desempeño cotidiano de las personas, ayudando, reemplazando, simulando las acciones o decisiones tomadas por individuos con ciertas características y conocimientos en particular. Como una de sus principales áreas de aplicación se puede destacar particularmente la medicina, ya que la Inteligencia Artificial logró su mayor impacto a través de los Sistemas Expertos y específicamente por medio de los sistemas de diagnóstico.

Muchas aplicaciones en la Inteligencia Artificial se han desarrollado con el fin de reforzar el cuidado de la salud y proporcionar mejores métodos para este objetivo particular. Como han expresado muchos estudios, las aplicaciones de la Inteligencia Artificial en el campo médico fueron desarrolladas para ayudar a los usuarios (particularmente médicos y pacientes), proporcionando diagnósticos y predicciones para evitar complicaciones en las enfermedades. Aunque estos sistemas están dotados con el conocimiento “humano”, nunca podrán reemplazar la experiencia humana, ya que como el mismo ser humano, estos sistemas exigen frecuentemente de una supervisión y una actualización de sus conocimientos, por consiguiente, los papeles del especialista (experto) médico y de los doctores continúan siendo importantes para asegurar la validez del sistema.

Los primeros estudios en sistemas médicos inteligentes tales como CASNET, MYCIN, PIP e Internist-I han mostrado la posibilidad de realizar diagnóstico en varios tipos de enfermedades. A continuación se describen las principales características de estos primeros sistemas basados en la Inteligencia Artificial como ayuda para el diagnóstico médico [42]:

- CASNET (Causal ASSociational Networks): se desarrolló a inicios de 1960, fue una herramienta general para construir Sistemas Expertos para el diagnóstico y tratamiento de enfermedades. La mayoría de las aplicaciones realizadas mediante el uso de CASNET se especializaron en el diagnóstico y tratamiento del glaucoma.
- MYCIN se desarrolló a inicios de 1970, está considerado el primer sistema experto de la historia y fue diseñado para diagnosticar ciertas infecciones microbianas, además de recomendar el tratamiento con la asignación respectiva de medicamento necesario para tratar la infección. MYCIN tenía varios módulos tales como: explicación, adquisición de conocimiento e interacción.
- PIP, que era una abreviación para el Present Illness Program, se desarrolló en 1970 para simular el comportamiento que debe seguir el especialista al momento de clasificar la historia médica de un paciente, diagnosticándolo o identificándolo dentro de algún cuadro de las enfermedades renales conocidas.
- El Internist-I se desarrollo a inicios de 1980, su función se centraba en la investigación de métodos heurísticos para tareas de diagnóstico diferencial usadas en la toma de decisiones clínicas (fue usado en diagnóstico de medicina interna).

Además de estos cuatro sistemas, que forman el conjunto de aplicaciones más destacadas de la Inteligencia Artificial dentro de la medicina, existen otras aplicaciones, que aunque menos importantes o aún no desarrolladas completamente, merecen ser nombradas en este documento ya que forman parte de la historia de la Inteligencia Artificial:

- ICHT (sistema basado en Inteligencia Artificial para el Cuidado de la Salud de los niños) se desarrolló para reducir la mortalidad de los niños sobre todo en las áreas rurales. El éxito de éste sistema radicaba en categorizar las insuficiencias más comunes dentro de la población infantil, tomando en consideración factores de riesgo importantes tales como el control del peso, inmunización, además de hitos en el crecimiento y la nutrición. ICHT utilizó Sistemas Expertos en el proceso de toma de los datos para la generación de la historia de los pacientes.
- HERMES (HEpathology Rule-based Medical Expert System) es un Sistema Experto para el pronóstico de las enfermedades crónicas más comunes.
- SETH es un Sistema Experto para el tratamiento del envenenamiento por el uso de drogas.
- PROVANES es un Sistema Experto híbrido para los pacientes críticos en anestesiología.

- ISS (Interactive STD Station) se desarrolló para el diagnóstico de las enfermedades de transmisión sexual.

El diagnóstico médico es una tarea que requiere de gran precisión, dadas las consecuencias que puede llevar consigo una decisión mal tomada. Por lo tanto dicha actividad es realizada por médicos con una cierta experiencia en el área en la que se este realizando el diagnóstico. La Inteligencia Artificial, como su mismo significado lo expresa, pretende emular la capacidad natural que posee el hombre en la toma de decisiones de cualquier tipo, imitando tanto su modo de aprendizaje como la manera en que basado en dicho conocimiento puede llegar a tomar decisiones. Para este propósito se han venido aplicando diferentes ramas de la Inteligencia Artificial dentro de las cuales se pueden destacar los Sistemas Expertos, las Redes Neuronales, la Minería de Datos o los Agentes Inteligentes de la Lógica Difusa.

Tal vez el mayor problema en el campo de la medicina es realizar el diagnóstico de una enfermedad, y uno de los inconvenientes más importantes en éste proceso es la subjetividad del especialista que lo realiza; este hecho se hace notar en particular en actividades de reconocimiento de patrones, donde la experiencia del profesional esta directamente relacionada con el diagnóstico final. Esto es debido al hecho de que el resultado no depende de una solución sistematizada sino de la interpretación de los síntomas del paciente.

Hay que destacar que la mayoría de médicos confrontan durante su formación la tarea de aprender a diagnosticar. En esta fase tienen que resolver el problema de deducir ciertas enfermedades o formular un tratamiento basado en observaciones o especificaciones de conocimiento. En términos generales, el diagnóstico médico es el proceso de identificar la enfermedad que está sufriendo un paciente, para de esta manera poder determinar cuál es la mejor forma para tratarla. Para el diagnóstico en general y el diagnóstico médico particularmente, la integración rápida y fácil de conocimiento que pueda reemplazar al antiguo es un factor fundamental, dado que el diagnóstico médico es un proceso muy complejo, que requiere la recopilación de los datos del paciente, un profundo entendimiento de la literatura médica alrededor del tema y muchos años de experiencia clínica; sin embargo un diagnóstico totalmente preciso no puede ser realizado sin antes considerar muchas alternativas. Como resultado de esta incertidumbre las decisiones tomadas por diferentes médicos en diferentes estados del proceso de diagnóstico no siempre son las mismas, porque cada caso particular tiene un proceso de decisión diferente para cada médico así se trate del diagnóstico del mismo tipo de enfermedad.

El diagnóstico médico apoyado mediante aplicaciones de Inteligencia Artificial se puede ver referenciado desde varias de sus principales técnicas, tales como:

Sistemas Expertos:

1. Diagnóstico basado en reglas.
2. Diagnóstico basado en probabilidades.

Lógica Difusa:

1. Diagnóstico basado en clasificación.

Redes neuronales:

1. Diagnóstico basado en entrenamiento.
2. Diagnóstico basado en reconocimiento.

Minería de datos aplicada:

1. Diagnóstico mediante el reconocimiento de patrones).

2.1.3.- Sistemas expertos

Cuando se dice que una persona es experta en algo quiere decir que esta persona tiene un conocimiento especializado sobre ese tema. En los Sistemas Expertos a este tipo de conocimiento se le llama conocimiento sobre el dominio.

Los Sistemas Expertos son un campo dentro de la Inteligencia Artificial. Estos sistemas también son conocidos como Sistemas Basados en Conocimiento y permiten la creación de máquinas que son capaces de razonar como el hombre, aunque en un espacio de conocimientos limitado y concreto (lo que se ha denominado dominio). En teoría pueden razonar siguiendo los pasos que seguiría un experto humano (médico, analista, empresario, etc.) para resolver un problema concreto, lo que ofrece diferentes posibilidades a la hora de resolver problemas y aprender.

El origen de estos sistemas se remonta a mediados de los años sesenta. Durante esta década los investigadores Alan Newell y Herbert Simon [43] desarrollaron un programa llamado GPS (General Problem Solver), que podía trabajar con criptoaritmética, con las torres de Hanoi y con otros problemas similares. Lo que no podía hacer era resolver problemas del mundo real, tales como un diagnóstico médico. Fue entonces cuando algunos investigadores decidieron limitar el problema a un dominio específico y simular el razonamiento de un experto. A partir de 1965, un equipo dirigido por Edward Feigenbaum, comenzó a desarrollar Sistemas Expertos utilizando bases de conocimiento definidas con mucho detalle. Dos años más tarde se

construyó DENDRAL, que se considera el primer Sistema Experto. El objetivo de este Sistema Experto era identificar estructuras químicas moleculares a partir de su análisis espectrográfico.

En la década de los setenta se desarrolló MYCIN, para la consulta y el diagnóstico de infecciones de la sangre. Este sistema introdujo nuevas características, como la utilización de conocimiento impreciso para razonar y la posibilidad de explicar el proceso de razonamiento. Lo más importante es que funcionaba correctamente, llegando a decisiones similares a las que llegaría un ser humano tras años de aprendizaje. Después apareció EMYCIN (MYCIN Esencial) con el que se construyó SACON, utilizado para estructuras de ingeniería, PUFF para estudiar la función pulmonar y GUIDON para elegir tratamientos terapéuticos.

En esa misma época se desarrollaron también los siguientes sistemas: HERSAY, que intentaba identificar la palabra hablada, y PROSPECTOR, utilizado para hallar yacimientos de minerales.

Posteriormente, en la década de los ochenta, se ponen de moda los Sistemas Expertos, surgiendo numerosas empresas de alta tecnología para la investigación en esta área de la Inteligencia Artificial, desarrollando incluso Sistemas Expertos para su comercialización. Se llega a la conclusión de que el éxito de un Sistema Experto depende casi exclusivamente de la calidad de su base de conocimiento. El inconveniente es que codificar la habilidad de un experto puede resultar difícil, largo y muy trabajoso.

Un ejemplo de Sistema Experto moderno es CASHVALUE, que evalúa proyectos de inversión y VATIA, que asesora acerca del impuesto sobre el valor añadido o IVA.

Por tanto, los Sistemas Expertos se pueden entender como programas de computación que almacenan conocimiento de expertos humanos en algún área específica, y luego permiten usar ese conocimiento para resolver problemas de una manera muy aproximada a la que el experto seguiría.

La principal forma de obtener resultados en un Sistema Experto es basándose en la generación de reglas. Estas reglas se obtienen del experto humano y representan el conocimiento obtenido mediante sentencias de la forma “Si [...] → Entonces [...]”. Las reglas pueden ser usadas para realizar operaciones entre los datos, inferir y encontrar la conclusión más apropiada. Por lo tanto se puede concluir que la potencia de un Sistema Experto se basa más que en una gran cantidad de conocimiento, en un algoritmo deductivo muy eficaz. La idea que se es la de automatizar la labor del experto, partiendo en ocasiones de información insuficiente o incompleta.

Diagnóstico basado en reglas

En la vida diaria se encuentran muchas situaciones complejas gobernadas por reglas deterministas, tales como sistemas de control de tráfico, sistemas de seguridad, transacciones bancarias.... Las reglas deterministas constituyen la metodología más sencilla de las utilizadas en Sistemas Expertos. La base de conocimiento contiene las variables y el conjunto de reglas que definen el problema, y el motor de inferencia obtiene las conclusiones aplicando la lógica clásica a estas reglas. La regla es una proposición lógica que relaciona dos o más objetos e incluye dos partes, la premisa y la conclusión [28]. Cada una de estas partes consiste en una expresión lógica con una o más afirmaciones objeto-valor conectadas mediante los operadores lógicos 'y', 'o', o 'no'. Una regla se escribe normalmente como "Si premisa" → "entonces conclusión".

Diagnóstico basado en probabilidades.

Este tipo de Sistemas Expertos surgen porque incertidumbre es lo común y no la excepción, ya que los datos pueden no ser conocidos con exactitud (un paciente puede no saber si ha tenido o no fiebre) y además porque el conocimiento no es determinista (un mismo conjunto de síntomas puede estar asociado a diferentes enfermedades). Se eligió la probabilidad como medida para tratar la incertidumbre, pero se encontraron problemas, debidos al uso incorrecto de las hipótesis de independencia, utilizadas para reducir la complejidad. Finalmente, tal y como afirmó Lindley, "la única descripción satisfactoria de la incertidumbre es la probabilidad. Esto quiere decir que toda la información incierta debe estar en forma de probabilidad, que varias incertidumbres deben ser combinadas usando las reglas de la probabilidad, y que el cálculo de probabilidades es adecuado para manejar situaciones que implican incertidumbre. Las descripciones alternativas de la incertidumbre son innecesarias" [60].

2.1.4.- Lógica difusa o borrosa

Es una técnica que usa la teoría matemática de grupos difusos, que simula el proceso de razonamiento humano, permitiendo al ordenador comportarse de una forma menos precisa y lógica que los computadores convencionales. Esta aproximación es usada particularmente porque la acción de tomar decisiones en el mundo real no siempre esta enmarcada en decisiones del tipo verdadero o falso, sino que estas decisiones pueden tener estados intermedios (términos de "tal vez").

La Lógica Difusa (o borrosa, del inglés *fuzzy*) es una rama de la Inteligencia Artificial que se basa en que "Todo es cuestión de grado", lo que permite manejar información de difícil especificación. Con la Lógica Difusa es posible desarrollar un sistema mediante reglas de "sentido común", las cuales se refieren a cantidades indefinidas. Las reglas involucradas en un sistema difuso, se pueden aprender con sistemas adaptativos que aprenden al "observar" como

trabajan las personas con los dispositivos reales, o pueden también ser formuladas por un experto humano. En general la Lógica Difusa se aplica tanto a sistemas de control como para modelar cualquier sistema continuo de ingeniería, física, biología o economía.

La Lógica Difusa se puede definir como un sistema matemático que modela funciones no lineales, que convierte unas entradas en salidas acordes con los planteamientos lógicos que usan el razonamiento aproximado. Se fundamenta en los denominados conjuntos difusos y un sistema de inferencia difuso basado en reglas de la forma “Si [...] \rightarrow Entonces [...]”, donde los valores lingüísticos de la premisa y el consecuente están definidos por conjuntos difusos.

En 1994, la teoría de la Lógica Difusa alcanzó su punto más alto, pero esta idea no era nueva, ya Aristóteles consideraba que existían ciertos grados de veracidad y falsedad. Platón consideró ya grados de pertenencia. David Hume describió que el núcleo de un concepto atrae conceptos similares, y además creía en la lógica del sentido común, el razonamiento basado en el conocimiento que se adquiere ordinariamente con las diferentes vivencias de cada uno. Por otro lado, Immanuel Kant consideraba que solo los matemáticos podían proveer definiciones claras, y muchos principios contradictorios no tenían solución [22].

Según el profesor Lotfi Zadeh, creador de la Lógica Difusa, “la lógica difusa constituye un sistema lógico que está dedicado a la formalización de modos de razonamiento que son aproximados y no exactos. En ese sentido la lógica borrosa es una extensión de los sistemas lógicos polivalentes, pero sus objetivos son bastante diferentes, tanto en espíritu como en esencia.”

2.1.5.- Redes neuronales

Son un modelo que simula una Red Neuronal Biológica. Este concepto se usa para implementar simulaciones de software en procesos paralelos masivos que envuelven el procesamiento de elementos interconectados en una arquitectura de red. Las neuronas artificiales reciben entradas que son equivalentes a los impulsos eléctricos que las dendritas de las neuronas biológicas reciben de otras neuronas. La salida de la neurona artificial equivale a la señal enviada fuera de la neurona biológica a través de su axón. Estas señales artificiales pueden ser cambiadas del mismo modo en que ocurren los cambios físicos en la sinapsis de la neurona biológica. Una Red Neuronal es “un nuevo sistema para el tratamiento de la información, cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso humano: la neurona” [9].

Las Redes Neuronales son unos sistemas que pueden simular ciertas características propias de los humanos, como la capacidad de memorizar y de asociar hechos. Si se observan los problemas que no pueden expresarse con un simple algoritmo se verá que tienen una característica en común, la experiencia. El hombre es capaz de resolver ciertos problemas

gracias a la experiencia acumulada, aproximándose al problema mediante la construcción de sistemas capaces de simular esta característica.

Diagnóstico basado en entrenamiento.

También se denominan redes de entrenamiento supervisado [17] y han sido los modelos más desarrollados de este tipo de diseños. Los datos para el entrenamiento se componen por varios pares de patrones de entrenamiento de entrada y de salida. El hecho de conocer la salida implica que el entrenamiento se beneficia de la supervisión de un maestro. En este modo de aprendizaje se muestran los patrones a la red y la salida deseada para esos patrones y se usa una fórmula matemática de minimización del error que ajuste los pesos para dar la salida más cercana posible a la salida deseada.

Diagnóstico basado en reconocimiento.

También se denominan redes de entrenamiento no supervisado [17] y utilizan datos de entrenamiento consistentes en sólo patrones de entrada, por lo tanto, la red es entrenada sin el beneficio de un maestro. La red aprende a adaptarse basada en las experiencias recogidas de los patrones de entrenamiento anteriores.

Este tipo de aprendizaje no necesita los patrones objetivos para salida, ya que el algoritmo y la regla de modificación de las conexiones producen patrones de salida consistentes. Cuando la red procesa patrones similares da la misma salida para ambos, es decir, clasifica los patrones en categorías parecidas, aunque esto es relativo ya que se debe decidir en qué grado deberán parecerse los patrones para asignarles la misma salida. Normalmente se usa el error cuadrático medio para determinar esta similitud, aunque puede haber otras opciones.

En este proceso de aprendizaje, la red mide cuánto se parece el patrón que le llega y los que tiene almacenados y en función del resultado los agrupa en una categoría u otra, aunque en un principio no se sepa qué salida corresponderá a cada tipo o grupo de patrones de entrada, ni que atributos usará para clasificarlos.

2.1.6.- Minería de datos aplicada

Es una técnica de Inteligencia Artificial usada para el descubrimiento de conocimiento en grandes conjuntos de datos, es decir, que se usa para descubrir información que a simple vista se encuentra oculta a los propósitos a los que se refiera el sistema.

La Minería de Datos es un proceso que extrae conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos

formatos. Las herramientas usadas predicen futuras tendencias y comportamientos, permitiendo la toma de decisiones.

La minería de datos es un proceso que invierte el método científico en el sentido de que, en el método científico, primero se formula la hipótesis y luego se diseña el experimento para coleccionar los datos que confirmen o rechacen esta hipótesis. En la Minería de datos, se coleccionan los datos y se espera que de ellos surjan hipótesis, luego se valida esa hipótesis basada en los datos, que será numéricamente significativa, pero experimentalmente inválida. Por este motivo esta técnica presenta un enfoque exploratorio y no confirmador.

Los métodos de Minería de Datos son aplicados para extraer conocimiento interesante y regular. Estos métodos pueden ser usados para entender los datos espaciales, descubrir relaciones entre datos espaciales y no espaciales, reorganizar los datos en bases de datos espaciales y determinar sus características generales de manera simple y concisa.

2.2.- Sistemas de clasificación automática de imágenes basados en contenido

2.3.1.- Introducción a los sistemas de clasificación de imágenes basados en contenido

Los sistemas de clasificación de imágenes basados en contenido, también denominados CBIR (*Content Based Image Retrieval*) son una de las aplicaciones de las técnicas de visión computerizadas con los que se intenta solventar el problema de clasificación de imágenes y su búsqueda en grandes bases de datos [19].

Que estos sistemas sean “basados en contenido” significa que en la búsqueda de las imágenes se analizará el contenido real de las mismas, en lugar de metainformación, como pueden ser palabras clave con las que se etiquetan las imágenes o descripciones propias de las mismas. El término “contenido” en este contexto se refiere al color, la forma, la textura o cualquier otra información que pueda derivarse de la imagen (o de la imagen después de haberla sometido a ciertos tratamientos como pueden ser filtros, transformadas...). Los resultados obtenidos en los sistemas que se basan en la búsqueda de metainformación (también denominados “basados en texto”) producen muchos resultados incorrectos o poco precisos, y es por este motivo por el que los sistemas basados en contenido son preferibles y los estudios sobre los mismo se incrementan cada día.

Los sistemas CBIR permiten, por tanto, encontrar de forma rápida y eficiente una imagen similar a la que se le proporciona dentro de una base de datos, que puede llegar a ser de tamaño más que considerable, mediante la comparación del “contenido” de esta imagen con

el “contenido” de las imágenes almacenadas en la base de datos y que han sido previamente etiquetadas o clasificadas.

El desarrollo de este sistema será específico para el contexto en el que se trabaje (en este caso el diagnóstico médico) ya que la base de datos estará organizada y codificada acorde a este contexto, y además porque dependiendo de las necesidades del problema concreto que se esté abordando, las características que mayor información proporcionan pueden variar. En los siguientes apartados se van a explicar tanto las distintas características que se pueden obtener de una imagen y la forma de extraerlas, así como los posibles tipos de codificación y bases de datos que existen, además de los distintos tipos de distancia que pueden usarse a la hora de comparar las características de las imágenes de consulta con las de referencia. Una vez vistas todas las posibilidades, en el siguiente apartado de este documento se van a detallar las que se han seleccionado para el diseño de este sistema de clasificación de imágenes en concreto.

De manera general estos sistemas de clasificación de imágenes se componen de dos módulos:

1.- Módulo de extracción de características: se extraen las características de las imágenes (previamente seleccionadas como las que más información proporcionan).

2.- Módulo de clasificación de imágenes: a partir de las características de una imagen de la que no se conoce la clasificación, el sistema la ubica dentro de una clase (o le proporciona una codificación).

Para la implementación completa del sistema de clasificación de imágenes se usan los dos módulos previamente descritos en dos fases distintas:

1.- Fase de entrenamiento: en esta fase se trabaja con imágenes cuya clase (o codificación) es conocida, para entrenar al sistema, de manera que se asocia esa clase a unos valores de las características de la imagen previamente extraídas. Al conocer la clasificación de la imagen, esta fase permite elegir las características que mayor información proporcionan y reajustar los parámetros del módulo de extracción de características para obtener los mejores resultados posibles.

2.- Fase de test: en esta fase se trabaja con imágenes cuya clase (o codificación) es desconocida, para evaluar el funcionamiento del sistema, que procederá a la clasificación de las mismas de manera que se pueda obtener la tasa de error para cuantificar la eficiencia del sistema.

2.3.2.- Características de las imágenes

De una imagen pueden obtenerse un gran número de características, que se van a dividir en dos grupos: características visuales de bajo nivel o características físicas, y características visuales de alto nivel o características lógicas basadas en el reconocimiento de objetos. A continuación se explican las diferentes características que forman estos dos grupos:

A. Características visuales de bajo nivel o características físicas [4]:

Describen características visuales básicas como son la forma, el color, la textura, las formas, el movimiento y las regiones, y permiten encontrar una conexión entre los píxeles contenidos en una imagen digital y lo que las personas recuerdan una vez que han observado una imagen. Estas características pueden expresarse cuantitativamente y se extraen de la imagen original. Se obtienen mediante procesamiento digital de señales.

COLOR: es la característica más básica del contenido visual de una imagen. El color es una percepción visual que se genera en el cerebro al interpretar las señales nerviosas que le envían los fotorreceptores de la retina del ojo y que a su vez interpretan y distinguen las distintas longitudes de onda que captan de la parte visible del espectro electromagnético.

TEXTURA: caracteriza las diferentes texturas o regiones de una imagen, la homogeneidad o heterogeneidad de las mismas y los histogramas de los bordes de estas regiones, por lo que no caracteriza píxeles de manera independiente, sino relaciones entre los mismos.

La textura es un ente visual que describe cierto orden estructural o tendencia en los elementos presentes en una imagen. La textura es, lógicamente, una ilusión óptica. La sensación de textura le da a una imagen una capacidad comunicativa extra. A través de ella, la persona que observa la imagen tiene la sensación de reconocer las sensaciones táctiles que experimentaría al tocar una determinada superficie.

La textura de una imagen se puede definir también como una región macroscópica estructurada que presenta propiedades locales constantes, que varían lentamente o aproximadamente periódicas. Este orden local viene de la presencia repetida y no casual de partes elementales que tienen dimensión similar [27].

La textura depende de la escala a la que se observa la imagen, ya que la textura de un píxel no está definida, sino que hay que estudiar la vecindad de un grupo de píxeles para poder obtener una descripción de la textura de la misma. Dependiendo del tamaño de vecindad elegido la textura puede ser diferente y tener diferentes características: uniformidad, densidad, tosquedad, aspereza, regularidad, linealidad, dirección, frecuencia, fase...

FORMA: los seres humanos son capaces de reconocer objetos simplemente viendo su forma, y por ello ésta es una característica muy importante. Esta información se extrae mediante segmentación (operación similar a la que realiza el sistema visual humano cuando ve un objeto). Se puede definir como la estructura expresiva plástica de un objeto, donde se asienta su identidad visual.

Esta característica permite obtener las diferentes regiones que se encuentran en una imagen, definiendo fronteras, estructuras... El estudio de la morfología de una imagen puede hacerse de varias formas, y se basa en la teoría de conjuntos, de manera que cada conjunto representa la forma de cada uno de los objetos que componen la imagen (los píxeles blancos se representan con un '1' y los negros con un '0' (o viceversa): morfología binaria, morfología de niveles de gris y morfología de imágenes policromáticas.

LOCALIZACIÓN: es una característica que alude a los distintos elementos que se encuentran dentro de una imagen y se utiliza para describir dichos elementos en el dominio espacial, identificando así su orientación dentro del conjunto.

B. Características visuales de alto nivel o características lógicas basadas en el reconocimiento de objetos.

Proporcionan información acerca de los objetos y eventos que van apareciendo en la escena y la relación que existe entre ellos.

Éstos descriptores, que proporcionan información sobre objetos concretos en la imagen, no son fáciles de extraer, menos aún en el caso de que se pretenda realizar una extracción automática de los mismos. Hay que tener en cuenta además que, al subdividir la imagen muchas veces la información sobre los diferentes objetos puede ser destruída, por lo que dentro de este tipo de características visuales se suelen usar diferentes métodos de segmentación de imágenes que permitan mantener los objetos que se buscan completos para poder reconocerlos.

C. Características invariantes.

Hay un tipo de características, que se denominan características invariantes que no varían al aplicarse ciertas transformaciones a las imágenes, como pueden ser rotaciones, traslaciones o interpolaciones. Son características que se tienen muy en cuenta a la hora de clasificar las imágenes ya que se pueden tener imágenes similares pero que hayan sufrido alguna de estas transformaciones.

2.3.3.- Extracción de características

Para el diseño e implementación del sistema de recuperación y clasificación de imágenes basado en contenido es necesario procesar e identificar las características de cada imagen a ser almacenada en la base de datos. Como se ha comentado anteriormente, estas características proporcionan información acerca del contenido de la imagen que permite distinguir una imagen de las demás (información única para cada imagen). Para recuperar la información necesaria para que el sistema funcione correctamente es necesario elegir una combinación de todas las características de las que se pueda disponer, que será diferente para cada contexto.

La primera aproximación que podría hacerse para comparar las características de las imágenes puede ser evaluar directamente la matriz de píxeles [21], y para ello ambas imágenes deben tener la misma dimensión. Esta solución no suele utilizarse por varios motivos; en primer lugar normalmente las imágenes no tienen la misma dimensión por lo que habría que hacer un re-escalado de las mismas. Por otro lado, para imágenes de gran tamaño puede ser una operación demasiado laboriosa debido al gran número de píxeles a comparar (las dimensiones de una imagen crecen linealmente con el re-escalado, pero el número de píxeles que la forman crece de manera cuadrática en función del área [11]). Pero el motivo más importante para no usar esta comparación de imágenes es que la información que da la comparación del valor de los píxeles de las mismas hace referencia únicamente al color, y no al resto de las características como pueden ser la forma, la textura o la orientación.

Por estos motivos se hace uso de otras herramientas que permiten realizar una comparación más general de las características de las imágenes. Para realizar esta comparación se suele hacer un re-escalado de las imágenes por varios motivos: en primer lugar para que ambas imágenes tengan el mismo tamaño y puedan compararse más fácilmente. En segundo lugar para reducir la complejidad de la matriz de píxeles de la imagen. A la versión escalada de la imagen se le llama "*down scaled image*", y con ella se consigue una muestra representativa de los píxeles originales (es común aplicar filtros sobre la imagen antes de hacer el escalado para resaltar detalles que puedan ser interesantes y así no perder esa información).

La extracción de características de las imágenes se va a hacer tanto de forma global extrayendo las características de la imagen completa, como de forma local segmentando la imagen y obteniendo las características de cada parte por separado. Hay varias opciones a la hora de dividir la imagen, ya que pueden usarse ventanas de tamaño fijo o se pueden buscar regiones previamente seleccionadas.

De esta manera, con las características globales y locales (en ambos casos, tanto físicas como lógicas) se obtendrá un conjunto de características que permiten obtener la máxima información posible para poder realizar la clasificación de las imágenes.

2.3.3.1.- Elementos de procesado del color

A) HISTOGRAMA

Es una representación gráfica de los valores de una variable en forma de barras, de manera que cada barra representa un valor de la variable y la altura de las mismas la frecuencia de aparición de ese valor concreto. En este caso cada barra representa un nivel de la escala de grises. Esta representación da información sobre el brillo y el contraste de la imagen, y puede ayudar a ajustar estos parámetros.

En el caso de imágenes a color el histograma proporciona la frecuencia de aparición de las distintas tonalidades de color; en caso de imágenes en blanco y negro da la información de las distintas tonalidades de la imagen dentro de la escala de grises. Para las imágenes en color tendrán que usarse tres histogramas, uno para cada color básico (que dependerá del modelo de representación de color que se haya escogido) y cada barra representa los colores entre el negro y el color al que pertenece el histograma; pero si se trabaja en escala de grises (como es este caso) se utilizará uno sólo que contenga los valores del blanco al negro.

Con esta herramienta se obtiene una descripción de imagen de manera global, de manera que si las barras son más altas (mayor frecuencia) en la zona de colores claros, la imagen será clara (brillante), y si son más altas en el lado de colores oscuros será oscura. De esta representación también se puede concluir que si los valores se concentran en un número pequeño de barras, es decir, que el histograma tiene un perfil estrecho, se estará trabajando con una imagen de bajo contraste, y si por el contrario es un histograma disperso el contraste será alto.

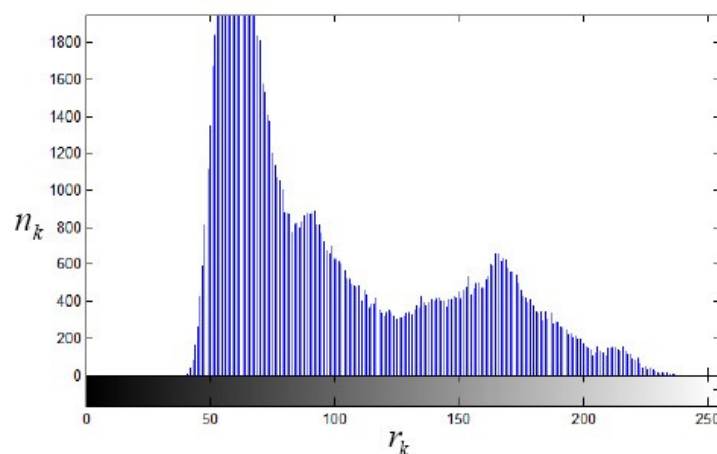


Figura 1. Ejemplo histograma

2.3.3.2.- Elementos de procesado de la textura

A) ESTADÍSTICOS

Son un conjunto de cinco estadísticos que se obtienen de forma sencilla y aportan información acerca de la difusión de la intensidad de la imagen y por tanto de su brillo. Estos cinco estadísticos se obtienen a partir de histograma de primer orden y son los expresados a continuación.

→ MEDIA: media aritmética, es decir, suma del valor de todos los píxeles que forman la imagen dividida por el número total de éstos. Da como resultado el centro de gravedad de los datos, que puede interpretarse como el nivel medio de intensidad que representa la imagen.

$$media = \frac{\sum_{i=1}^N \sum_{j=1}^M pixel(i, j)}{N \cdot M}$$

→ MEDIANA: es el valor de la intensidad del píxel que deja el mismo número de píxeles a un lado y a otro.

→ VARIANZA: es el promedio de las desviaciones cuadráticas con respecto a la media, es decir, que da una idea de cómo se distribuyen los píxeles de la imagen con respecto al valor medio de la intensidad, es decir, de la dispersión de los valores representados en el histograma (anchura del mismo).

$$varianza = \sigma^2 = \frac{\sum_{i=1}^N \sum_{j=1}^M (pixel(i, j) - media)^2}{N \cdot M}$$

→ COEFICIENTE DE ASIMETRÍA: este parámetro es un momento de tercer orden que da una idea de la asimetría de la distribución de los valores de intensidad de los píxeles respecto a la media. Si se obtiene un valor positivo la distribución está sesgada hacia la izquierda, si es negativo hacia la derecha y si es cero la distribución será simétrica.

$$asimetria = \frac{\sum_{i=1}^N \sum_{j=1}^M (pixel(i, j) - media)^3}{N \cdot M} \left[\sqrt{\sum_{i=1}^N \sum_{j=1}^M (pixel(i, j)^2 - media^2)} \right]^3$$

→ COEFICIENTE DE KURTOSIS: este parámetro es un momento de cuarto orden que da una idea sobre el aplanamiento de la distribución respecto a una distribución normal. Si se obtiene un valor mayor que 3 la distribución será más alta que la distribución normal, si es menor que 3 será una distribución más aplanada que la distribución normal, y si es 3 será igual a una distribución normal.

$$kurtosis = \frac{\sum_{i=1}^N \sum_{j=1}^M (pixel(i, j) - media)^4}{N \cdot M \left[\sqrt{\sum_{i=1}^N \sum_{j=1}^M (pixel(i, j)^2 - media^2)} \right]^4}$$

B) DIMENSIÓN FRACTAL

Este es un parámetro que da información sobre la rugosidad y la autosimilaridad de la imagen. El concepto de autosimilaridad da una idea de si las diferentes partes de la imagen tienen una estructura similar a la de la imagen global. Permite reconocer imágenes irregulares mediante unos objetos que se denominan fractales matemáticos, que son objetos semigeométricos cuyas estructuras básicas, fragmentadas o irregulares, se pueden repetir a diferentes escalas [7].

Los fractales tienen las siguientes características:

- Son demasiado irregulares para describirlos de forma geométrica.
- Poseen detalle a cualquier escala de observación.
- Son autosimilares.
- Se definen mediante algoritmos recursivos.

Algunos objetos de la naturaleza, como las nubes, las montañas, las líneas costeras, los copos de nieve o el sistema circulatorio (y otras estructuras del cuerpo humano) son fractales naturales. El fractal ofrece una representación aproximada de los objetos ya que el mundo natural tiene límites que sólo podrían sobrepasarse con fractales ideales.

La dimensión fractal se calcula a partir de una serie de fórmulas que tienen que ver con el recuento de las bolas necesarias para recubrir el conjunto o con el de cajas de una cuadrícula que contienen parte del conjunto, cuando las dimensiones de unas y otras tienden a cero (*Box-counting*). Este proceso ignora las irregularidades de tamaño menor y estudia cómo varía el número de celdas que contiene el objeto cuando la dimensión de las mismas tiende a cero. Se hace esta aproximación porque los objetos de la naturaleza no cumplen la totalidad de las características que tiene un cumplir un fractal para ser considerado como tal. Se puede

medir la dimensión fractal de objetos reales, y con estas medidas se pueden comparar objetos del mundo real con fractales generados por algoritmos matemáticos. Esta dimensión se expresa en términos del mínimo número de bolas de radio ϵ necesarias para recubrir el conjunto $N(\epsilon)$, de la siguiente manera [6]:

$$DF = \lim_{\epsilon \rightarrow 0} \frac{\ln N(\epsilon)}{\ln(1/\epsilon)}$$

O también puede definirse en función del recuento del número de cajas N_n de una cuadrícula de anchura $1/2^n$ que intersecan al conjunto, de la siguiente manera [6]:

$$DF = \lim_{n \rightarrow \infty} \frac{\ln N_n}{\ln(2^n)}$$

Ambas definiciones son equivalentes.

El análisis de la textura de una imagen basándose en fractales fue introducido por Pentland [46]. En este estudio se demuestra la correlación que existe entre la relación fractal de una textura y su tosquedad (*coarseness*). Además, gracias a la propiedad de autosimilaridad se tiene que la dimensión fractal de una imagen es independiente de su escala.

C) CARACTERÍSTICAS DE GABOR (FILTROS DE GABOR)

Los filtros de Gabor son filtros lineales cuya respuesta al impulso es una función senoidal multiplicada por una función gaussiana [23]. Al introducir esta envolvente gaussiana se consigue que las funciones de Gabor estén localizadas tanto en el dominio espacial como en el de la frecuencia (no como las funciones senoidales que sólo están localizadas en el dominio de la frecuencia), lo que deja una serie de funciones que permiten representar señales en ambos dominios indistintamente.

Al hacer la transformada de Fourier de un filtro de Gabor se obtiene una gaussiana centrada en la frecuencia de la función senoidal.

El filtrado de una imagen con filtros de Gabor está relacionado con los procesos en la corteza visual del ser vivo y se ha empleado en el procesamiento digital de imágenes ya que se ha comprobado que es eficiente a la hora de realizar segmentación de texturas, compresión de imágenes, etc.

La forma del filtro de Gabor está dada por la siguiente expresión:

$$G(x, y, f_0, \sigma_x, \sigma_y, \Theta_k) = \exp\left(-\frac{1}{2}\left(\frac{x_{\Theta_k}^2}{\sigma_x^2} + \frac{y_{\Theta_k}^2}{\sigma_y^2}\right)\right) \cdot \exp(2\pi i \cdot f_0 \cdot x_{\Theta_k}) \text{ con } k = 1, 2, \dots, m$$

Donde:

- $x_{\Theta_k} = x \cdot \cos(\Theta_k) + y \cdot \sin(\Theta_k)$
- $y_{\Theta_k} = -x \cdot \sin(\Theta_k) + y \cdot \cos(\Theta_k)$
- f_0 : frecuencia de la onda senoidal plana
- m: cantidad de orientaciones
- Θ_k : k-ésima orientación del filtro de Gabor
- σ_x desviación típica de la envolvente Gaussiana en el eje x
- σ_y desviación típica de la envolvente Gaussiana en el eje y

Para aplicar el filtro de Gabor se necesitan tres parámetros previamente especificados:

- Frecuencia de la onda senoidal plana en una dirección dada (puede darse también la longitud de onda que está directamente relacionada con la frecuencia).
- Orientación del filtro (se cuantizan todas las direcciones para cubrir el espacio $[0 \pi]$).
- Las desviaciones típicas en ambos ejes, σ_x y σ_y .

D) CARACTERÍSTICAS DE LAWS (FILTROS DE LAWS)

Estas características dan información sobre la energía de la estructura de una imagen [23], y se calculan a partir de una convolución de filtros con la imagen a estudiar y luego realizando una operación de enventanado no lineal.

El banco de filtros con el que se convoluciona la imagen se obtiene de la siguiente manera:

Se tienen cinco filtros de longitud cinco en 1D, que se emplean para la convolución en 2D, definidos por los siguientes coeficientes:

$$L5 = [1 \ 4 \ 6 \ 4 \ 1]$$

$$E5 = [-1 \ -2 \ 0 \ 2 \ 1]$$

$$S5 = [-1 \ 0 \ 2 \ 0 \ -1]$$

$$W5 = [-1 \ 2 \ 0 \ -2 \ 1]$$

$$R5 = [1 \ -4 \ 6 \ -4 \ 1]$$

- L5 (*Level filter*): realiza un promedio ponderado en 1D.
- E4 (*Edge filter*): detector de bordes
- S5 (*Spot filter*): detector de impulsos
- W5 (*Wave filter*): detector de ondas
- R5 (*Ripple filter*): detector de rizados

Con estos cinco filtros se pueden generar 25 máscaras diferentes de filtrado bidimensional, de la siguiente manera:

L5L5	L5E5	L5S5	L5W5	L5R5
E5L5	E5E5	E5S5	E5W5	E5R5
S5L5	S5E5	S5S5	S5W5	S5R5
W5L5	W5E5	W5S5	W5W5	W5R5
R5L5	R5E5	R5S5	R5W5	R5R5

Tabla 1. Máscaras de filtros bidireccionales de Laws

A continuación se van a explicar los pasos que hay que seguir para obtener el banco de filtros:

Paso 1: obtención de las máscaras y filtrado de la imagen: se realiza la convolución de la imagen de tamaño NxM con cada una de las 25 máscaras enumeradas en la tabla anterior. Se obtienen 25 imágenes filtradas de tamaño NxM en escala de grises, que serán la base utilizada para el análisis de la textura de la imagen original.

Paso 2: enventanado: se reemplaza el valor de cada píxel de cada una de las 25 imágenes filtradas por otro valor que dependa del valor de los píxeles vecinos. Para ello se usan ventanas de tamaño 15x15 (píxeles) y se aplica la siguiente expresión:

$$new(x, y) = \sum_{i=-7}^7 \sum_{j=-7}^7 old(x+i, y+j)$$

Se obtienen 25 imágenes TEM (*Texture Energy Measure*):

L5L5T	L5E5T	L5S5T	L5W5T	L5R5T
E5L5T	E5E5T	E5S5T	E5W5T	E5R5T
S5L5T	S5E5T	S5S5T	S5W5T	S5R5T
W5L5T	W5E5T	W5S5T	W5W5T	W5R5T
R5L5T	R5E5T	R5S5T	R5W5T	R5R5T

Tabla 2. Imágenes después del enventanado

Paso 3: normalización de las características para el contraste: se normalizan las 25 imágenes de la tabla anterior respecto a la media del filtro L5L5 (que es la única que no es nula) y se descarta la imagen L5L5T.

Paso 4: combinación de características similares: se mezclan las características que se obtienen con las máscaras traspuestas. Estas son las nuevas características:

$$\begin{aligned}E5L5TR &= E5L5T + L5E5T \\S5L5TR &= S5L5T + L5S5T \\W5L5TR &= W5L5T + L5W5T \\R5L5TR &= R5L5T + L5R5T \\S5E5TR &= S5E5T + E5S5T \\W5E5TR &= W5E5T + E5W5T \\R5E5TR &= R5E5T + E5R5T \\W5S5TR &= W5S5T + S5W5T \\R5S5TR &= R5S5T + S5R5T \\R5W5TR &= R5W5T + W5R5T \\E5E5TR &= E5E5T * 2 \\S5S5TR &= S5S5T * 2 \\W5W5TR &= W5W5T * 2 \\R5R5TR &= R5R5T * 2\end{aligned}$$

Este conjunto está formado por las 14 texturas que se pueden obtener por imagen, que son invariantes a las rotaciones (cada píxel se representa por 14 rasgos de textura). Para hacer la comparación de imágenes y después su clasificación lo que se hace es obtener la energía de cada una de las 14 texturas de las imágenes

E) CARACTERÍSTICAS DE TAMURA

Los filtros de Tamura proporcionan seis características que son útiles para el análisis de la textura de una imagen [58]. Estas seis características son: tosquedad (*coarseness*), contraste (*contrast*), direccionalidad (*directionality*), línea de semejanza (*line-likeness*), regularidad (*regularity*) y rugosidad (*roughness*). Las tres primeras tienen una gran semejanza con la percepción humana de las imágenes, de manera que podemos tenerlas muy en cuenta a la hora de identificar y clasificar las imágenes. Son características asociadas a cada píxel de la imagen, y si se unen se puede reconstruir la imagen aunque sin información sobre el color.

Tosquedad (*coarseness*): proporciona información acerca del tamaño de las texturas de la imagen. Si las texturas siguen un macro-patrón y un micro-patrón, siempre se considerará el patrón más grande. La forma de hacerlo es aplicando operadores de distinto tamaño tal y como se explica a continuación:

Paso 1. Para cada píxel (x, y) se calcula el promedio sobre una vecindad definida que debe ser de tamaño potencia de 2:

$$A(x, y) = \frac{1}{2^{2k}} \sum_{i=1}^{2^k} \sum_{j=1}^{2^k} pixel(x - 2^{k-1} + i, y - 2^{k-1})$$

Paso 2. Para cada punto (x, y) se calcula la diferencia entre las vecindades que no se superpone a lados opuestos del punto en las direcciones horizontal y vertical:

$$D_k^h(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)|$$

$$D_k^v(x, y) = |A_k(x, y + 2^{k-1}) - A_k(x, y - 2^{k-1})|$$

Paso 3. Para cada punto (x, y) se selecciona el valor con la mayor diferencia:

$$S(x, y) = \max_{d=h,v} \{E_k^d(x, y)\}, k = \{1, \dots, 5\}$$

Este valor se asigna a cada píxel, obteniendo la tosquedad local de la imagen, lo cual permite calcular una distribución de tosquedad en la imagen, es decir, un histograma que permita medir los tamaños de las texturas.

Paso 4. Finalmente, se toma la media de S(x, y) a lo largo de la imagen completa:

$$F_{tosq} = \frac{1}{m \times n} \cdot i \sum_i^m \sum_j^n S_{best}(i, j)$$

Siendo “m” la altura de la imagen y “n” la anchura.

Contraste (*contrast*): es una característica que viene influida por los niveles de grises, la polarización en los niveles de blanco y negro, la definición de los bordes y el periodo de repetición de los patrones:

$$F_{cont} = \frac{\sigma}{\alpha_4^z} \quad \text{con} \quad \alpha_4^z = \frac{\mu_4}{\sigma_4}$$

$$\mu_4 = \frac{1}{XY} \sum_{x=1}^X \sum_{y=1}^Y (pixel(x, y) - \mu)^4$$

Donde “ μ_4 ” es el cuarto momento alrededor de la media y “ σ_4^{2z} ” es la varianza de los valores de grises en la vecindad y “z” ha sido configurado experimentalmente como ¼.

Direccionalidad (*directionality*): define la orientación de la textura:

$$\Delta_H = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \Delta_V = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

$$|\Delta G| = \frac{\Delta_H + \Delta_V}{2}$$

Para cada posición (x, y) se calcula el ángulo de orientación:

$$\theta = \frac{\pi}{2} + \tan^{-1} \left(\frac{\Delta_H(x, y)}{\Delta_V(x, y)} \right)$$

El valor de la orientación será un valor comprendido entre 0 y pi, medido en el sentido de las agujas del reloj para que la dirección horizontal sea el cero. El histograma H_D se obtiene cuantificando θ y contando el número de puntos con una magnitud $|\Delta G|$ por encima de un umbral "t".

Para mediar cuantitativamente la direccionalidad del histograma lo que se hace es medir la agudeza de las crestas del mismo. La aproximación que se realiza es sumar los segundos momentos alrededor de cada máximo de un valle a otro valle:

$$F_{dir} = 1 - r \cdot n_p \cdot \sum_p \sum_{\phi \in w_p} (\phi - \phi_p)^2 \cdot H_D(\phi)$$

Donde " n_p " es el número de picos, " ϕ_p " es la posición del pico p-ésimo de H_D y " w_p " es el rango del pico p-ésimo entre valles.

Línea de semejanza (line-likeness): cuando la dirección de un píxel y la de sus vecinos es muy cercana, todos ellos se representan mediante una única línea. De manera que se construye una matriz de dirección de concurrencia cuyo elemento $P_{Dd}(i, j)$ se define como una frecuencia relativa con la que dos celdas vecinas se encuentran separadas una distancia "d" a lo largo de la dirección de borde que ocurre en la imagen, una con dirección "i" y la otra con dirección "j".

$$F_{lin} = \sum_i \sum_j P_{Dd}(i, j) \cdot \cos(i - j) \cdot \frac{2\pi}{n} \bigg/ \sum_i \sum_j P_{Dd}(i, j)$$

Donde P_{Dd} es la matriz de concurrencia local de tamaño nxn de los puntos a una distancia "d".

Regularidad (regularity): describe los patrones repetitivos de forma matemática, aunque no resulta sencillo analizar la regularidad de los objetos naturales, ya que resulta complicado

dar un grado de irregularidad sin tener datos sobre el tamaño o la forma del elemento. Se asume que si cualquier característica de textura varía a lo largo de toda la imagen, dicha imagen es irregular. Igualmente, se puede dividir la imagen en varias subimágenes y considerar la variación de cada una de las características en cada subimagen. Como medida de regularidad, se puede considerar la suma de la variación de las cuatro características anteriormente comentadas:

$$F_{reg} = 1 - r \cdot (\sigma_{tosq} + \sigma_{cont} + \sigma_{dir} + \sigma_{lin})$$

Donde “r” es un factor de normalización y “ σ_x ” es la desviación estándar de cada una de las características implicadas.

Rugosidad (roughness): está ligada a la tosquedad y el contraste, por lo que, de forma aproximada, podría decirse que:

$$F_{reg} = F_{tosq} + F_{cont}$$

F) MATRIZ DE CO-OCURRENCIA

A la matriz de co-ocurrencia también se le llama *Gray-Level Co-Occurrence Matrix (GLCM)* y es un instrumento muy útil para el análisis textural de imágenes y la discriminación de patrones [52]. Esta matriz aporta información espacial en forma de posición relativa entre distintos niveles de intensidad dentro de la textura. Una vez obtenida esta matriz se caracteriza la información de la textura por medio de sus momentos estadísticos de segundo orden.

Para construir la matriz hay que definir una serie de parámetros:

→ Condición: regla que debe cumplir un determinado píxel para poder ser cuantificado con relación a un vecino. Suele tomarse la ubicación del píxel de determinada intensidad que se encuentra en la dirección del vecino.

→ Dirección: orientación espacial en la cual se evalúa la condición (valores típicos 0°, 45°, 90°, 135°).

→ Distancia: número de píxeles a los cuales se encuentran el par de píxeles evaluados en la condición (valores típicos 1, 3 y 5 píxeles).

En la matriz aparece la probabilidad de que los dos píxeles a una distancia “d” y en una orientación “ ϕ ” tengan el mismo nivel de gris.

$$GLCM_d^\phi(i1, i2) = p(T(u, v) = i1, T(u \pm df(\phi)) = i2)$$

Los estadísticos de segundo orden que se calculan son los siguientes:

1.- Segundo momento angular, Energía, Uniformidad

Se obtiene mediante la siguiente expresión:

$$energia = \sum_{i=0}^{255} \sum_{j=0}^{255} GLCM(i, j)^2$$

Esta propiedad da una idea de la suavidad de la textura. Cuanto más suave es la textura mayor valor toma la energía. Si la intensidad es constante, la energía toma valor 1, mientras que si se tiene una distribución uniforme de parejas, será menor que 1.

Cuanto más semejantes sean los valores de la matriz, es decir, cuanto más dispersa sea la diagonal (por ser una matriz simétrica), el valor de la uniformidad será menor; pero si en la diagonal se dan mayores picos de intensidad el descriptor será máximo.

2.- Contraste

Se obtiene mediante la siguiente expresión:

$$contraste = \sum_{i=1}^{255} \sum_{j=1}^{255} (i - j)^2 \cdot GLCM(i, j)$$

Esta propiedad da una idea acerca de las variaciones bruscas de niveles de gris en la imagen.

El valor del contraste aumentará, si existen más elementos de la matriz alejados de la diagonal principal. Hay que destacar que estos elementos de la diagonal principal tienen un efecto despreciable ya que no proporcionan información acerca del contraste.

Una textura de características suaves y uniformes tendrá un contraste bajo, pero una con aspecto rugoso e irregular lo tendrá alto.

3.- Entropía

Se obtiene mediante la siguiente expresión:

$$entropia = \sum_{i=1}^{255} \sum_{j=1}^{255} GLCM(i, j) \cdot \log_2 GLCM(i, j)$$

Esta propiedad da una idea de la aleatoriedad de la matriz. Si los elementos de la matriz son similares, el descriptor aumenta su valor (siendo máximo si todos los valores de la matriz son iguales).

A medida que aumenta la uniformidad de la textura de la imagen el valor de la entropía irá descendiendo.

4.- Momento diferencial inverso, Homogeneidad

Se obtiene a partir de la siguiente expresión:

$$\text{homogeneidad} = \sum_{i=1}^{255} \sum_{j=1}^{255} \frac{GLCM(i, j)}{1 + (i - j)^2}$$

Esta propiedad da una idea de la regularidad local de la textura. El valor del descriptor aumenta a medida que disminuye $(i - j)$.

Cuanto menor sea el contraste mayor será la homogeneidad debido a la relación inversa con $(i - j)^2$.

5.- Correlación

Se obtiene mediante la siguiente expresión:

$$\text{correlacion} = \sum_{i=1}^{255} \sum_{j=1}^{255} \frac{(i - \mu_x) \cdot (j - \mu_y)}{\sigma_x \cdot \sigma_y} \cdot GLCM(i, j)$$

Es una medida de probabilidad que mide la relación entre las diferentes intensidades de los niveles de gris. Su valor está ligado a sus medidas estadísticas.

El valor del descriptor aumentará a medida que aumente la distancia de los elementos a la media; de manera similar disminuye cuando la varianza es baja, es decir, que los elementos de la matriz no se encuentran muy alejados de la diagonal principal.

2.3.3.3.- Elementos de procesado de la forma

Para analizar la forma de una imagen se usan elementos de procesado morfológico, que al ser desplazados alrededor de la misma dan información sobre la posición relativa del elemento con respecto al plano y al fondo de la imagen. Una vez se tiene el elemento de procesado morfológico se sitúa el centro del mismo en cada píxel de la imagen original y se

aplica una operación morfológica sobre todos los píxeles que quedan situados bajo el elemento. Hay distintos tipos de elementos de procesamiento morfológico:

- **Anti-extensivos:** el resultado de la operación morfológica está contenido en el conjunto original.
- **Idempotentes:** el resultado siempre es el mismo independientemente de las veces que se aplique.
- **Extensivos:** el resultado obtenido de la operación morfológica contiene al conjunto original.

Las diferentes operaciones morfológicas que se pueden realizar con estos elementos son [43]:

- **Dilatación:** se reemplaza cada píxel blanco de la imagen original por una réplica del elemento estructurante. Es una operación extensiva y permite rellenar agujeros y bahías de tamaño igual o menor que el elemento estructurante. Usado con imágenes en escala de grises, la dilatación se efectúa tomando el máximo de una serie de sumas. Puede ser usado para implementar el operador de 'máxima vecindad' con la forma de la vecindad dada en el elemento estructurante.

Cabe destacar que, técnicamente, el grupo de píxeles de la imagen original sobre el que se realiza la operación morfológica no es una imagen en sí misma, sino un conjunto, y el resultado de la dilatación tampoco es una imagen sino otro conjunto. La salida de la dilatación es el conjunto de puntos barridos por el centro del elemento estructural mientras algún punto del mismo coincide con alguno de la imagen original, de manera que la dilatación puede interpretarse como el resultado de reemplazar cada píxel blanco de la imagen original por una réplica del elemento estructurante.

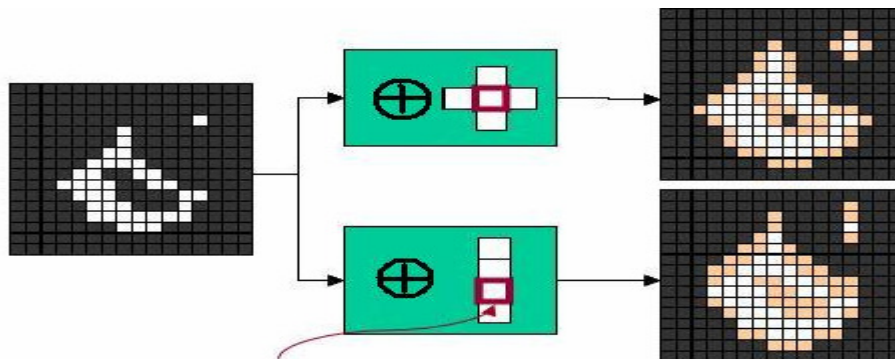


Figura 2. Operación morfológica de dilatación

En el ejemplo, los píxeles que cambian de negro a blanco por efecto de la dilatación aparecen ligeramente coloreados.

- **Erosión:** se elimina cada píxel de la imagen original en los que el elemento estructurante no cabe. Es una operación anti-extensiva y permite eliminar pequeñas islas y protuberancias de tamaño menor o igual que el elemento estructurante. Usado en imágenes en escala de grises, la erosión se efectúa tomando el mínimo de una serie de diferencias. Puede ser usado para implementar el operador de 'mínima vecindad' con la forma de la 'vecindad' dada por el elemento estructurante.

La salida de la erosión es el conjunto de puntos barridos por el centro del elemento estructural mientras se cumpla que todos los puntos del mismo estaban contenidos en la imagen original. Elimina grupos de píxeles donde el elemento estructural no cabe.

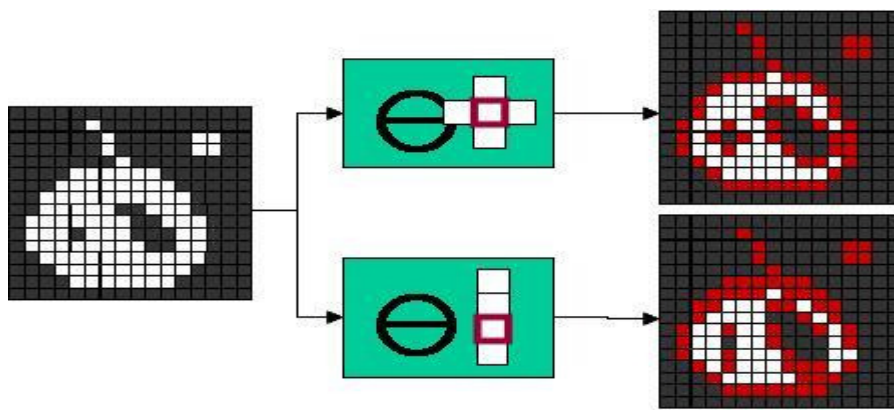


Figura 3. Operación morfológica de erosión

En el ejemplo, los píxeles que cambian de blanco a negro por efecto de la erosión aparecen en rojo.

Se pueden combinar ambas operaciones, dando lugar a lo que se conoce como filtros morfológicos, que también se utilizan para el análisis de la forma de los objetos dentro de las imágenes. Existen dos posibles filtros:

- **Filtros de apertura:** combina un operador de erosión y otro de dilatación con el mismo elemento estructurante. Es una operación anti-extensiva e idempotente y permite alisar contornos, eliminar protuberancias y separar objetos en puntos estrechos. Puede obtenerse desplazando el elemento estructural por el interior del conjunto y eliminando aquellas zonas por las que no pueda “pasar”.

Esta operación tiene un inconveniente, y es que tanto la erosión como la dilatación son operaciones no son invertibles, por lo que la imagen original no se podrá recuperar.

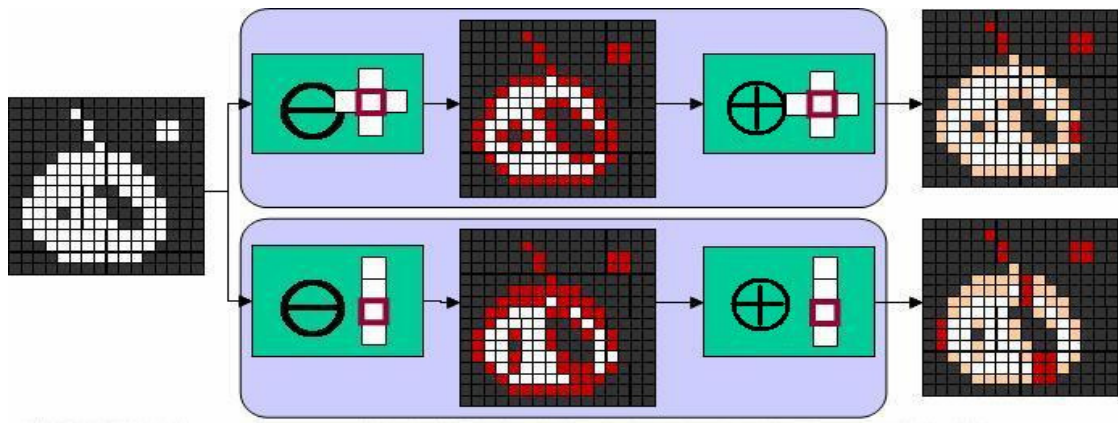


Figura 4. Filtro de apertura

• **Filtros de cierre:** combina un operador de dilatación y otro de erosión con el mismo elemento estructural. Es una operación extensiva e idempotente y permite alisar porciones de contornos, fusionar grietas estrechas, rellenar vacíos y agujeros pequeños, eliminar entrantes pequeños y conectar objetos vecinos. Puede obtenerse desplazando el elemento estructural por el exterior del conjunto y eliminando aquellas zonas por las que no pueda “pasar”.

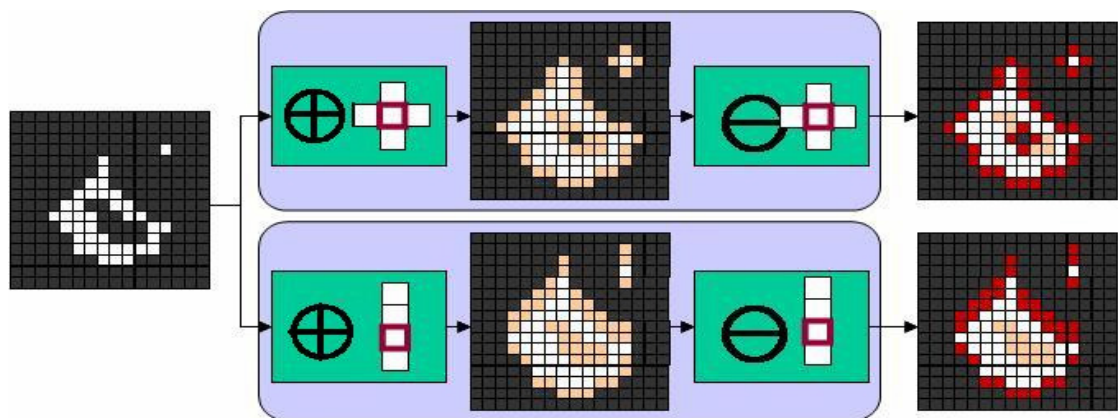


Figura 5. Filtro de cierre

2.3.3.4.- Elementos de procesamiento de características invariantes

Hay varias formas de construir características invariantes [5]:

→ Normalización: se consigue una representación similar para imágenes que pertenezcan a la misma clase. De esta manera, se busca un miembro representativo para cada clase de objeto y se definen contornos normalizados a las posibles transformaciones que puede sufrir la imagen. Es un procedimiento muy sensible al ruido y no aporta información acerca de la localización espacial.

Una posible forma de normalización de una imagen es calculando el centro de gravedad de los niveles de gris de la imagen y transformándolo en el centro de gravedad de la misma (se usa para normalizar respecto a la traslación de un nivel de gris).

→ Aproximación diferencial: se consigue una descripción de los objetos a partir de un pequeño conjunto de diferenciales invariantes que contengan información esencial, de manera que las características diferenciales seleccionadas deben permanecer constantes frente a pequeñas transformaciones infinitesimales.

Para obtener las características diferenciales mediante este método es necesaria la segmentación de la imagen, ya que lo que se usa son los contornos de los objetos procesados.

$$\frac{\delta F(g(\lambda) \cdot I)}{\delta \lambda_i} \Big|_{\lambda=0} = 0$$

Siendo: “F”: conjunto de características invariantes que se obtienen.

“g”: transformación aplicada a la imagen.

“I”: imagen sobre la que se aplica la transformación.

“λ”: parámetros que se varían.

→ Aproximación integral: se hace la integración de todas las imágenes transformadas (se consideran todas las posibles transformaciones), de manera que la integral es invariante frente a las transformaciones consideradas.

$$F(I) = \frac{1}{|G|} \int f(gI) dg$$

Uno de los problemas de la aproximación integral es que la cantidad de datos que es proporcionada para comparar imágenes no es muy grande, por lo que se recurre a los siguientes procedimientos:

A) HISTOGRAMA DE CARACTERÍSTICAS INVARIANTES

Se sustituyen una o varias integrales por histogramas:

$$histograma_F(X) = \frac{1}{R} \sum_{r=1}^R f(gI)$$

B) VECTORES DE CARACTERÍSTICAS INVARIANTES

Se calculan varias características invariantes utilizando diferentes funciones f_i de manera que se obtiene un vector $V = (v_1, v_2, \dots, v_N)$ de características invariantes con $v_n = F_n(X)$ y f_n son funciones del conjunto descrito anteriormente.

C) CARACTERÍSTICAS INVARIANTES DE FOURIER MELLIN

La amplitud del espectro que se obtiene al aplicar la transformada de Fourier a una imagen es invariante frente a la traslación. Si se usan coordenadas polares se pueden crear características invariantes respecto a la rotación y el escalado además de la traslación [5]. De esta manera se tiene que la amplitud del espectro es invariante respecto a la traslación, varía inversamente respecto al escalado y es variante respecto a la rotación.

Si se usan las coordenadas polares y una escala logarítmica para el eje radial, las imágenes escaladas y rotadas se convierten en una representación de manera que la amplitud de la misma es invariante con respecto al escalado, la rotación y la traslación.

2.3.3.5.- Elementos de procesamiento de las características lógicas

La segmentación permite reconocer objetos concretos dentro de una imagen. Puede hacerse con una ventana de tamaño limitado o buscando objetos con una forma determinada. Para buscar objetos que tengan una forma determinada dentro de una imagen se hace a través de una serie de características concretas de esos objetos.

El grado de segmentación de una imagen depende del objeto que se esté buscando o del grado de detalle que se necesite en la aplicación, de manera que terminará una vez se hayan encontrado los objetos deseados.

Hay diferentes algoritmos que permiten segmentar imágenes que se basan en las propiedades básicas de los valores de la escala de grises de las imágenes: discontinuidad y similitud [47]. La discontinuidad permite segmentar la imagen basándose en los cambios abruptos de nivel de gris, permitiendo localizar puntos aislados, líneas y bordes de una imagen. La similitud permite segmentar la imagen basándose en diferentes niveles de umbral, permitiendo localizar regiones de crecimiento del nivel de gris, y corte y pegado de regiones. El valor de este umbral se puede determinar a partir del histograma de la imagen, diferenciando las diferentes regiones por la frecuencia de aparición de los distintos niveles de gris.

2.3.3.6.- Transformaciones

Las transformaciones permiten trabajar con una misma imagen en diferentes dominios, como pueden ser el dominio temporal y el dominio de la frecuencia. Esto se hace debido a que

se pueden encontrar características que ayuden a clasificar las imágenes en ambos dominios y que proporcionen información relevante que ayude en la tarea de clasificación de imágenes. Generalmente estas transformaciones suelen tener una transformada inversa de manera que aunque se realice la operación se pueda volver a recuperar la imagen original.

Las operaciones realizadas sobre la imagen en el dominio espacial y en el dominio transformado son totalmente diferentes, por eso permite acceder a características totalmente diferentes que si únicamente trabajáramos en uno de los dominios. En algunos casos las características obtenidas en el dominio transformado también se pueden obtener en el dominio original, pero con operaciones mucho más complejas desde el punto de vista del cálculo.

De forma general, las transformadas permiten extraer características nuevas o de forma más sencilla, compresión de la imagen original, filtrado de la imagen original y reconocimiento de patrones. Todas estas acciones permiten restaurar imágenes, realizar operaciones como convolución y deconvolución, realzado de imágenes, detección de rectas/curvas, segmentación de imágenes, búsqueda de patrones, rotaciones, análisis espectral en tiempo real...

Una transformada de una imagen es una función que permite representarla como una combinación lineal de ciertas funciones base, de manera que:

$$imagen_transformada = c \cdot \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) \cdot imagen_original(x,y)$$

Las variables “x” e “y” están relacionadas con el dominio del tiempo, y “u” y “v” con el dominio de la frecuencia.

Las posibles transformaciones habituales que podemos aplicar sobre una imagen son las siguientes.

A) TRANSFORMADA DISCRETA DE FOURIER (DFT)

Los vectores base de la transformada discreta de Fourier son funciones exponenciales complejas que siguen la siguiente expresión [1]:

$$a_{u,v}(x,y) = \exp(ju \frac{2\pi}{M} x + jv \frac{2\pi}{N} y) = \cos(u \frac{2\pi}{M} x + v \frac{2\pi}{N} y) + j \sin(u \frac{2\pi}{M} x + v \frac{2\pi}{N} y)$$

De esta manera, la expresión de la transformada discreta de Fourier es la siguiente:

$$DFT\{f(x, y)\} = F_{DFT}(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cdot \exp(-ju \frac{2\pi}{M} x - jv \frac{2\pi}{N} y)$$

Las aplicaciones de esta transformada son muy variadas, desde filtros basados en máscaras, filtros lineales para eliminar ruido o interferencias periódicas, hasta realzar los bordes de una imagen, etc.

El principal problema de la transformada discreta de Fourier es que, al transformar la imagen al dominio de la frecuencia, se pierde información espacial. Como se comentó anteriormente, puede ser de gran utilidad, para mantener esta información, aplicarla mediante ventanas.

B) TRANSFORMADA DISCRETA DEL COSENO (DCT)

Es una transformada basada en la transformada de Fourier discreta pero utilizando únicamente números reales (en lugar de trabajar con exponenciales complejas, lo hace sólo con cosenos). Expresa una secuencia finita de varios puntos como resultado de la suma de distintas señales senoidales (con distintas frecuencias y amplitudes) [18].

Los vectores base de la transformada discreta del coseno son funciones coseno muestreadas y por ello el resultado siempre es real. Además, estos vectores dependen únicamente del orden de la transformada, y no de las propiedades estadísticas del conjunto de datos de entrada.

La transformada discreta del coseno unidimensional sigue la siguiente expresión:

$$DCT\{f(x)\} = F_{DCT}(u) = \alpha(u) \cdot \sum_{x=0}^{N-1} f(x) \cdot \cos\left(\frac{(2x+1) \cdot u \cdot \pi}{2N}\right) \quad \text{para } u = 0 : N-1$$

$$\text{Donde: } \alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & \text{para } u = 0 \\ \sqrt{\frac{2}{N}} & \text{para } u = 1 : N-1 \end{cases}$$

De manera similar, se puede definir la transformada discreta del coseno bidimensional:

$$DCT\{f(x, y)\} = F_{DCT}(u, v) = \alpha(u) \cdot \alpha(v) \cdot \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cdot \cos\left(\frac{(2x+1) \cdot u \cdot \pi}{2M}\right) \cdot \cos\left(\frac{(2y+1) \cdot v \cdot \pi}{2N}\right) \quad \text{para } u = 0 : M-1; v = 0 : N-1$$

Esta transformación de la imagen es muy útil para la compresión de la misma, por varios motivos:

- La DCT tiene una buena capacidad de compactación de la energía al dominio transformado, es decir, que la transformada de coseno discreta consigue concentrar la mayor parte de la información en pocos coeficientes transformados.
- La transformación es independiente de los datos. El algoritmo aplicado no varía con los datos que recibe, como si sucede en otros algoritmos de compresión.
- Tiene una interpretación frecuencial de los componentes transformados. La capacidad de interpretar los coeficientes en el punto de vista frecuencial permite aprovechar al máximo la capacidad de compresión.

Puede tener otras aplicaciones, como el reconocimiento de patrones o el filtrado, pero sin duda la compresión de imágenes es la más importante. La hace de manera que supone los coeficientes transformados de orden superior nulos. Lo primero que se debe hacer es obtener bloques de $N \times N$ píxeles de la imagen y aplicar a cada uno de esos bloques la expresión de la transformada discreta del coseno bidimensional. A la esta transformada se le pasan matrices $N \times N$ y el resultado serán matrices $N \times N$ en el dominio frecuencial. Una vez obtenida la matriz se le aplicara una reordenación en zig-zag. Lo que se hace es trabajar con las frecuencias espaciales que hay en la imagen, de manera que las altas frecuencias corresponderán a niveles de detalles altos, y las bajas, a niveles bajos. Conforme se recorre la matriz los primeros valores que se encuentran pertenecen al espacio de frecuencias más bajo (nivel de detalle más bajo), y los valores finales del recorrido se corresponden con las altas frecuencias (nivel de detalle más alto).

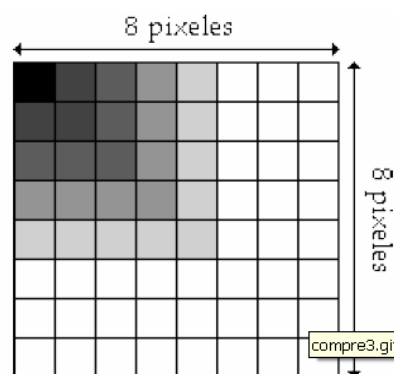


Figura 6. Niveles relevantes de los coeficientes de la transformada discreta del coseno

Que los distintos coeficientes de la transformada sean incorrelados es muy importante para compresión, ya que, el posterior tratamiento de cada coeficiente se puede realizar de forma independiente, sin pérdida de eficiencia de compresión.

150	80	40	14	4	2	1	0
92	75	36	10	6	1	0	0
52	38	26	8	7	4	0	0
12	8	6	4	2	1	0	0
4	3	2	0	0	0	0	0
2	2	1	1	0	0	0	0
1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Figura 7. Matriz obtenida de la transformada discreta del coseno

150	80	40	14	4	2	1	0
92	75	36	10	6	1	0	0
52	38	26	8	7	4	0	0
12	8	6	4	2	1	0	0
4	3	2	0	0	0	0	0
2	2	1	1	0	0	0	0
1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Figura 8. Reordenación en zig-zag de la transformada discreta del coseno

C) TRANSFORMADA DE KARHUNEN - LOÈVE (KLT)

En la teoría de procesos estocásticos, el Teorema de Karhunen-Loève es una representación de un proceso estocástico como una combinación lineal infinita de funciones ortogonales. Esta representación es análoga a la representación en series de Fourier de una función definida en un intervalo acotado de números reales. A diferencia de una serie de Fourier, en la cual los coeficientes son números reales y la base de expansión está compuesta por funciones senoidales, los coeficientes del teorema de Karhunen-Loève son variables aleatorias y la base de expansión depende del proceso. De hecho, la base de funciones ortogonales que se usa para la representación queda determinada por la función covarianza del proceso [67].

Esta transformada está basada en las propiedades estadísticas de la imagen y además no cumple la propiedad de separabilidad (en módulo y fase).

Las principales aplicaciones de esta transformada son la compresión y la rotación de imágenes. Generalmente se aplica a imágenes monocromas, y las imágenes en color son digitalizadas como grupos de tres imágenes, separando así los componentes rojo, verde y azul.

Esta transformada es óptima para codificación, ya que concentra el máximo de energía en el menor número de bases, pero sólo se puede utilizar pre-entrenándola para el grupo de

imágenes de interés, por lo que se suele utilizar más la transformada discreta del coseno, sin embargo es muy utilizada como extractor de características previo a la decisión. Sus principales utilidades son el reconocimiento de letras y de caras.

D) TRANSFORMADA DE HOUGH (HGT)

La transformada de Hough es un algoritmo empleado en reconocimiento de patrones y caracterización de imágenes, que permite encontrar ciertas formas dentro de una imagen, como líneas, círculos, etc. Su funcionamiento consiste en que para cada punto que se desea averiguar si es parte de una línea se aplica una operación dentro de cierto rango, con lo que se averiguan las posibles líneas de las que puede ser parte el punto. Esto se continúa para todos los puntos en la imagen, al final se determina qué líneas fueron las que más puntos posibles tuvieron y esas son las líneas en la imagen [67].

La transformada emplea una representación paramétrica de formas geométricas. Una recta, por ejemplo se representa por un módulo ϕ (phi) (perpendicular a la recta y que pasa por el origen (0,0) y un ángulo ρ (rho) (formado por el módulo y el eje positivo de las x's). La expresión de la recta mediante este algoritmo es la siguiente:

$$x \cos(\rho) + y \sin(\rho) = \phi$$

La ventaja de este método es que evita singularidades, como por ejemplo rectas de pendiente infinita. Si se representa ϕ y ρ en un plano cartesiano, una recta queda determinada mediante un punto con coordenadas ($\phi(\text{recta}), \rho(\text{recta})$), mientras que un punto, se representa como una función senoidal. Si por ejemplo se tienen dos puntos, habrá dos senoides desfasadas α grados dependiendo de las coordenadas de los puntos. Dichas senoides se irán cruzando cada 180° . La interpretación geométrica de este hecho, es que la función seno de cada punto, representa las infinitas rectas que pasan por cada punto, cuando dos puntos comparten la misma recta, sus representaciones senoidales se cruzan, se obtiene un punto. Cada vez que se da media vuelta ($\rho=180^\circ$) se vuelve a repetir la misma recta, por lo que se volverá a obtener otro punto, que de hecho es la misma recta.

E) TRANSFORMADA DE WAVELET DISCRETA (DWT)

La transformada de wavelet es un tipo especial de transformada de Fourier que representa una señal en términos de versiones trasladadas y dilatadas de una onda finita (denominada óndula madre) [40].

La transformada Wavelet Discreta se utiliza mucho en el campo de procesamiento de imagen gracias a su flexibilidad en la representación de señales no estacionarias y a sus posibilidades de adaptabilidad a la visión humana.

Gracias a su relación con el análisis multiresolución y descomposición en subbandas es muy útil para comprimir información, es decir, para extraer pocos parámetros representativos de la imagen. Este método de representación ofrece ortogonalidad y la posibilidad de representación tiempo-frecuencia.

El funcionamiento de la transformada se basa en que mediante filtros se descompone recurrentemente una señal de entrada en dos subseñales de menor resolución consideradas como aproximación y detalle. Sus expresiones son las siguientes:

$$aproximacion = S_{i+1}(n) = \sum_k g(k) \cdot S_i(2n - k)$$

$$detalle = W_{i+1}(n) = \sum_k h(k) \cdot S_i(2n - k)$$

El diagrama de bloques es el mostrado en la siguiente figura.

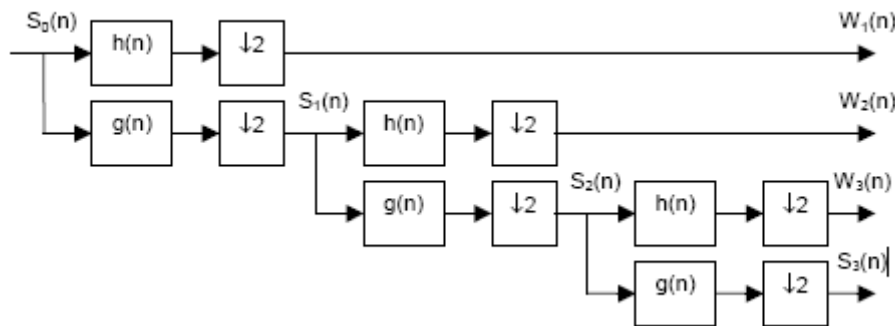


Figura 9. Diagrama de bloques de la transformada discreta de Wavelet

Al realizar la transformada a una imagen se obtiene una matriz de coeficientes, conocidos como coeficientes wavelet, que se dividen en cuatro tipos distintos:

- Aproximaciones.
- Detalles horizontales.
- Detalles verticales.
- Detalles diagonales.

La aproximación contiene la mayor parte de la energía de la imagen, es decir, la información más importante, mientras que los detalles tienen valores próximos a cero.

La elección de las wavelets analizadoras juega un papel muy importante en los resultados finales. Entre las características más importantes a tener en cuenta se encuentran: soporte compacto, simetría, etc. Las wavelets biortogonales son las más eficientes para un

posterior proceso de compresión, y en particular, aquellas con pocos coeficientes, ya que el coste de obtención de los coeficientes se incrementa con su número.

Generalmente, la energía de las imágenes se concentra en las frecuencias bajas. Una imagen tiene un espectro que se reduce con el incremento de las frecuencias. Estas propiedades de las imágenes quedan reflejadas en la transformada wavelet discreta de la imagen. Los niveles más bajos de compresión se corresponden con las bandas de alta frecuencia. En particular, el primer nivel representa la banda de más alta frecuencia y el nivel más fino de resolución. A la inversa, el último nivel (n) de descomposición corresponde con la banda de frecuencia más baja y la resolución más tosca. Así, al desplazarse de los niveles más altos de compresión a los más bajos, o sea, de baja resolución a alta resolución, se observa una disminución de la energía contenida en las subbandas recorridas.

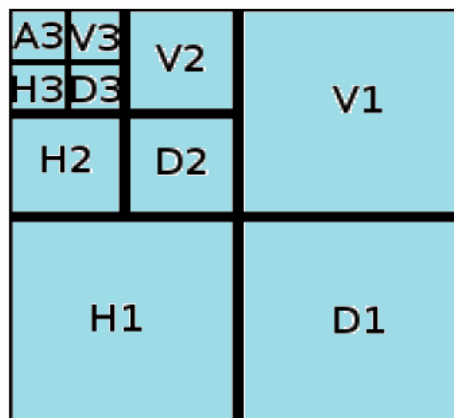


Figura 10. Esquema de coeficientes de Wavelet

Si los coeficientes wavelet obtenidos por medio de la transformada wavelet discreta para un nivel concreto poseen pequeñas magnitudes (valores próximos a cero), se espera que esos coeficientes estén en los primeros niveles de descomposición. El aumento del nivel de descomposición produce unos coeficientes con mayores magnitudes.

2.3.4.- Clasificación y codificación de imágenes

Una vez se han obtenido todas las características de la imagen, y se han estudiado de manera que se ha obtenido el conjunto de todas ellas que nos proporciona mayor información para nuestra aplicación en concreto (tanto de forma global como de forma local), se han de llevar a cabo las tareas de clasificación y codificación de las mismas para introducirlas en la base de datos y poder trabajar con ellas, bien para realizar consultas, bien para clasificar nuevas imágenes que necesitemos introducir en ella.

La codificación se realiza mediante el código estándar IRMA (*Image Retrieval in Medical Applications*), y se basa en un sistema ortogonal de cuatro ejes con tres o cuatro posiciones. Cada uno de estos ejes da un tipo de información independiente de la que da el resto, de forma jerárquica. La información que ofrece esta codificación es:

Técnica: tipo de imagen y forma de adquisición.

De dirección: orientación de la imagen con respecto al cuerpo.

Anatómica: región del cuerpo examinada.

Biológica: sistema biológico examinado.

El diagrama de bloques del proceso de clasificación y codificación de imágenes es el siguiente:

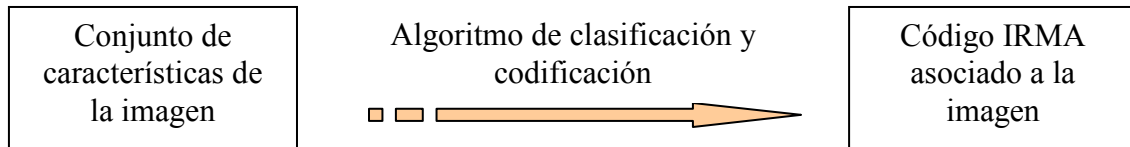


Figura 11. Diagrama de bloques del proceso de clasificación y codificación de imágenes

El algoritmo del vecino más cercano, k-NN, es un algoritmo de clasificación supervisada que se basa en el aprendizaje y realiza una estimación basada en un conjunto de entrenamiento de imágenes prototipo que le sirve para estimar la función de densidad $F(x | C_j)$ de las predictoras x por cada clase C_j .

Es un método de clasificación no paramétrico que estima el valor de la función de densidad de probabilidad o directamente la probabilidad a posteriori de que un elemento x pertenezca a una clase C_j a partir de la información proporcionada por el conjunto de prototipos. En este proceso de aprendizaje no se hace ninguna suposición acerca de la distribución de las variables predictoras.

La base del algoritmo es que es probable que las propiedades de un punto de entrada x sean similares a las de los puntos cercanos a éste [53]. El tamaño de la vecindad depende de cada implementación y hay que tener en cuenta varios aspectos: generalmente valores grandes de k reducen el efecto de ruido en la clasificación, pero crean límites entre clases parecidas.

Si el valor de k es muy pequeño, no contendrá ningún punto de los datos. Si es muy grande, puede incluir a todos los puntos de los datos, dando lugar a una estimación muy poco precisa. Hay que llegar a un compromiso entre: un valor lo suficientemente grande como para que los errores a la hora de clasificar las imágenes sean mínimos (estimación con significado), pero lo suficientemente pequeño para que los puntos muestrales estén lo suficientemente cerca como para pertenecer a la misma imagen.

Para los casos en que se produzca empate entre dos o más clases se tiene que definir una regla para el desempate, por ejemplo: seleccionar la clase que contenga al vecino más próximo o seleccionar la clase con distancia media menor. Para evitar que se produzcan empates, se suele tomar k impar.

Para llevar a cabo el proceso de clasificación y codificación de las imágenes se tiene un conjunto de n vectores (número de imágenes cuya clase se conoce previamente) que tienen m componentes (número de características obtenidas de cada imagen). Este conjunto de características de cada imagen, junto con la clase a la que pertenece cada una de ellas, se guarda en una matriz, que tiene la estructura siguiente.

$$\begin{bmatrix} c_{11} & c_{21} & c_{31} & c_{41} & c_{51} & c_{61} & c_{71} & c_{81} & c_{91} & C_1 \\ c_{12} & c_{22} & c_{32} & c_{42} & c_{52} & c_{62} & c_{72} & c_{82} & c_{92} & C_2 \\ c_{13} & c_{23} & c_{33} & c_{43} & c_{53} & c_{63} & c_{73} & c_{83} & c_{93} & C_3 \\ c_{14} & c_{24} & c_{34} & c_{44} & c_{54} & c_{64} & c_{74} & c_{84} & c_{94} & C_4 \\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{65} & c_{75} & c_{85} & c_{95} & C_5 \end{bmatrix}$$

Figura 12. Matriz de características

Las imágenes a las cuales pertenecen las características almacenadas en la matriz anterior forman lo que se denomina conjunto de imágenes de entrenamiento, y son las que constituirán la base de datos del sistema, ya que están previamente clasificadas, y permitirán, mediante la comparación de los vectores de características, clasificar imágenes cuya clase no es conocida previamente.

Y por otro lado se tiene otro conjunto de imágenes que no están previamente clasificadas, pero de las que se tiene el mismo conjunto de características que de las anteriores, por tanto, a partir de este conjunto de características, hay que determinar la clase de cada una de ellas haciendo uso del algoritmo de clasificación anteriormente descrito. Una vez clasificadas estas imágenes pasan a formar parte de la base de datos del sistema, de manera que también se añaden a la matriz y se utilizan para la clasificación de nuevas imágenes.

El algoritmo lo que hace es calcular la distancia que existe entre el vector de características de la imagen que se quiere clasificar y cada uno de los vectores de características de cada una de las imágenes previamente clasificadas que se tienen, de manera que esta imagen pertenecerá a la misma clase que aquella cuya distancia entre ambos vectores sea mínima. Podemos usar diferentes tipos distintos de distancia, pero típicamente se usan:

→ **Distancia euclídea:** es la distancia "ordinaria" entre dos puntos de un espacio euclídeo que se deduce a partir del teorema de Pitágoras. La distancia euclídea presenta la siguiente expresión:

$$d_{euclídea}(\vec{c1}, \vec{c2}) = \sqrt{(c_{11} - c_{12})^2 + (c_{21} - c_{22})^2 + (c_{31} - c_{32})^2 + \dots + (c_{n1} - c_{n2})^2}$$

→ **Distancia de Mahalanobis:** es una medida de distancia cuya utilidad radica en que es una forma de determinar la similitud entre dos variables aleatorias multidimensionales. Se

diferencia de la distancia euclídea en que tiene en cuenta la correlación entre las variables aleatorias. La distancia de Mahalanobis presenta la siguiente expresión [44]:

$$d_{euclídea}(\vec{c1}, \vec{c2}) = \sqrt{\left(\frac{c_{11} - c_{12}}{\sigma_1}\right)^2 + \left(\frac{c_{21} - c_{22}}{\sigma_2}\right)^2 + \left(\frac{c_{31} - c_{32}}{\sigma_3}\right)^2 + \dots + \left(\frac{c_{n1} - c_{n2}}{\sigma_n}\right)^2}$$

Donde σ es la desviación típica de los valores.

El principal problema de este método está en la carga computacional asociada al cálculo de las distancias a cada imagen del conjunto de entrenamiento.

Una vez que se tienen los k vecinos más cercanos a un vector de características de una imagen que se quiere clasificar y codificar, hay que decidir que código se le asignamos de entre los posibles (se tendrán k códigos distintos como mucho, a no ser que alguno se repita dentro de los k vecinos más cercanos, en cuyo caso el número de códigos posibles será menor que k). Para asignar un código a una imagen se siguen los siguientes pasos:

1.- Obtención de los k vecinos más cercanos a la imagen de test y su correspondiente código IRMA.

Una vez determinados cuáles son los k vecinos más cercanos de la imagen de test, se almacena el código IRMA de cada uno de ellos.

2.- Determinar la probabilidad de cada uno de los códigos.

Es la probabilidad de ocurrencia de ese código dentro de los k vecinos más cercanos, de manera que sigue la siguiente expresión:

$$probabilidad_codigo = \frac{n^\circ\text{ ocurrencias}}{k}$$

Se ha usado la probabilidad de ocurrencia del código completo en lugar de dividirlo en los diferentes ejes.

Llegados a este punto el resultado será el código que mayor porcentaje presente. En caso de empate deberíamos definir una regla que permitiera elegir uno u otro.

2.3.5.- Base de datos

La base de datos que se va a usar en este proyecto se denomina IRMA (Image Retrieval in Medical Applications) y es una base de datos que cuenta con 11.000 imágenes médicas divididas en 116 clases, de personas anónimas de distintas edades y con patologías diferentes. Estas 11.000 imágenes están divididas en dos conjuntos, un conjunto de entrenamiento de 10.000 imágenes y otro de test de 1.000 imágenes (codificadas con el código IRMA).

Las clases describen partes del cuerpo (las 8 clases generales son abdomen, cráneo, pecho, extremidades inferiores, extremidades superiores, pelvis, senos y columna vertebral). El tamaño de todas ellas es variable, pero siempre menor o igual que el cuadro de 512x512 píxeles, manteniendo la relación de aspecto original [37].

Este código permite clasificar y recuperar imágenes médicas de una manera más eficiente que otros estándares de comunicación actuales como *Digital Imaging and Communications in Medicine*, DICOM [68]. Aunque almacenar descripciones de forma explícita en las imágenes DICOM y los informes paralelos pueden suponer una ventaja a la hora de recuperar imágenes de forma selectiva, la terminología encontrada en DICOM no es lo suficientemente detallada para ordenar las entradas del sistema e incluso para obtener los resultados de búsqueda deseados.

El código IRMA intenta hacer frente a esos problemas para una recuperación óptima de las imágenes médicas, detallando: (a) el tipo de la misma junto a sus parámetros técnicos, (b) la orientación de la imagen con respecto al cuerpo, (c) la región del cuerpo examinada, y (d) el sistema biológico investigado [38].

La ventaja que presenta este código es que puede extenderse fácilmente introduciendo nuevos caracteres en ciertas posiciones del mismo, por ejemplo, si se introducen nuevas modalidades de adquisición de imágenes. Además, no presenta ningún tipo de ambigüedad en cuanto a un posible solapamiento de códigos de un mismo nivel.

El sistema de codificación IRMA [37] consiste en cuatro ejes con tres o cuatro posiciones, en cada una de las cuales aparecen $\{0, \dots, 9, a, \dots, z\}$, donde "0" denota "no especificado" para determinar el fin de un camino a lo largo de un eje:

A. Código técnico (TTTT)

Describe el tipo de imagen y se compone de cuatro dígitos.

1.- Fuente de adquisición de la imagen (por ejemplo: 1 "rayos-X", 2 "ultrasonidos", 3 "resonancia magnética", 4 "imagen óptica").

2.- Detalles más específicos de la fuente de adquisición de la imagen (por ejemplo: 11 "radiografía en una película a través de una proyección plana", 13 "angiografía", 14 "tomografía computerizada",...).

3.- Técnica de adquisición empleada (por ejemplo: 111 "digital", 112 "analógica", 113 "estereométrica", 114 "estereografía",...).

4.- Subtécnicas de adquisición (por ejemplo: 1111 "tomografía", 1112 "alta energía", 1113 "baja energía"). Se proporciona un listado completo del código IRMA T en el anexo A. Código IRMA.

B. Código direccional (DDD)

Describe la posición de la imagen respecto al cuerpo y se compone de tres dígitos:

1.- Orientación común (por ejemplo: 1 "coronal", 2 "sagital", 3 "transversal", 4 "otro").

2.- Especificación más detallada de la orientación común (por ejemplo: 11 "posterior-anterior (PA)", 12 "anteriorposterior (AP)").

3.- Momento en el que se ha llevado a cabo la captura de la imagen (por ejemplo: 111 "inspiración", 112 "expiración").

C. Código anatómico (AAA)

Describe la región del cuerpo examinada y se compone de tres dígitos:

1.- Regiones mayores (por ejemplo: 1 "cuerpo completo", 2 "cabeza", 3 "espina", 4 "extremidad superior",...).

2 y 3.- La región principal, a su vez, se detalla con dos dígitos de forma jerárquica (por ejemplo: 3 "espina", 31 "espina cervical", 311 "cubiles").

D. Código biológico (BBB)

Describe el sistema biológico examinado y se compone de tres dígitos:

1.- Diez posibles sistemas (por ejemplo: 1 "sistema cerebroespinal", 2 "sistema cardiovascular", 3 "sistema respiratorio", 4 "sistema gastrointestinal",...).

2 y 3.- Detalles de los sistemas que ayudan a identificar el órgano en cuestión (por ejemplo: 1 "sistema cerebroespinal", 11 "sistema nervioso central", 111 "cerebro").

El código IRMA resultante será un conjunto de 13 caracteres (IRMA: TTTT-DDD-AAABBB). Se trata de un sistema ortogonal, donde cada uno de los ejes ofrece resultados independientes del resto y lo hace de forma jerárquica.

A continuación se van a describir brevemente otras bases de datos utilizadas por algunos de los sistemas CBIR más conocidos.

A) Corel

Corel [69] posee una amplia base de datos de diferentes escenas. El número de imágenes oscila entre las 30.000 del sistema BlobWorld, y las 200.000 del sistema SIMPLYcity. FIRE utiliza 45.000 imágenes de 354X256 o 256X354 píxeles, dependiendo de la orientación.

Las imágenes de esta base de datos están descritas por texto por diferentes personas, por lo que la calidad de las anotaciones varía mucho entre las distintas imágenes (en alguna incluso las anotaciones aparecen con faltas de ortografía). La base de datos contiene en total 184.988 palabras que describen imágenes, de las cuales son diferentes 13.811.

B) MPEG-7

La base de datos de MPEG-7 posee 2.343 imágenes. Al igual que en la base de datos de Corel, las imágenes tienen un tamaño de 354 X 256 o 256 X 354 píxeles dependiendo de la orientación. La aplicación SIMBA utiliza esta base de datos.

C) COIL (Columbia Object Image Library)

COIL [69] es una base de datos para el reconocimiento de objetos, por lo que todas sus imágenes representan objetos, tanto en color como en blanco y negro. COIL posee dos bases de datos diferenciadas:

- COIL-20: Imágenes en escala de gris de 20 objetos diferentes.
- COIL-100: 100 imágenes en color de 72 objetos diferentes de 128X128 píxeles.

2.4.- *ImageCLEF: Anotación automática de imágenes médicas*

La tarea de anotación médica de ImageCLEF se realiza desde el año 2005 y forma parte de Cross Language Evaluation Forum (CLEF), un evento para la evaluación de sistemas de recuperación de información multilingüe y multimedia que se viene realizando desde el año

2000 [62]. ImageCLEF es uno de los congresos del área más importantes que existe en la actualidad, donde los mejores investigadores están presentes. La anotación o clasificación de imágenes médicas puede ser un importante primer paso para la búsqueda de imágenes en una base de datos. Para llevar a cabo esta tarea, la organización proporciona una base de datos con 11.000 imágenes clasificadas para llevar a cabo el entrenamiento del sistema de clasificación diseñado por cada grupo participante. Igualmente proporciona 1.000 imágenes de las cuales se desconoce su clasificación y que deben ser etiquetadas por los participantes. Una vez clasificadas, los resultados se envían y la organización procede a su evaluación. Terminado el plazo de entrega, se publica un ranking de los grupos participantes.

Las imágenes proporcionadas están en formato .PNG y la colección pertenece al grupo IRMA de la Aachen University of Technology (RWTH) en Alemania. En este proyecto cooperan el Department of Diagnostic Radiology, el Department of Medical Informatics, Division of Medical Image Processing y el Chair of Computer Science VI de la Aachen University of Technology con el objetivo de desarrollar y aplicar métodos de alto nivel para la recuperación de imágenes médicas basado en el contenido de las mismas, para la ayuda al diagnóstico. La colección incluye imágenes que corresponden a edades y géneros diferentes, así como adquisición de las mismas en posiciones diversas y múltiples posibles patologías. Por consiguiente, la calidad de la imagen varía significativamente.

El objetivo de los trabajos para ImageCLEF 2005 era determinar a cuál de las 57 posibles clases, identificadas arbitrariamente como números, pertenecía una imagen de test cuya clase era desconocida, por lo que cada participante mediante el sistema diseñado con las imágenes de entrenamiento, debía enviar las posibles clases de las imágenes de test. La organización de ImageCLEF, conocedora de las clases correctas de las imágenes de test, calculaba cuáles eran correctas y cuáles incorrectas, dando el “*error rate*” (%).

En ImageCLEF 2006 para valorar los resultados de anotación de los participantes, se calculaba la tasa de error de los envíos realizados, al igual que en el año anterior. El objetivo de los trabajos para ImageCLEF 2006 era determinar a cuál de las 116 posibles clases, identificadas también como números, pertenecía una imagen de test cuya clase era desconocida.

En ImageCLEF 2007, las 11.000 imágenes de entrenamiento están clasificadas según el código IRMA. Basándose en este código, se definen 116 categorías, cada una de las cuales se corresponde con un código IRMA diferente. A continuación se muestra el trabajo realizado por algunos de los participantes a lo largo de estos tres años en la tarea de anotación de imágenes médicas clasificados de acuerdo al tipo de técnica utilizada.

- **Grupo CYU_IM**

Pei-Cheng Cheng y Wei-Pang Yang [13] proponen un sistema de clasificación de imágenes médicas empleando el método del vecino más cercano (k-NN) con la ayuda de un vector de representación relativo que indica la relación espacial entre píxeles vecinos.

- **Grupo rwth_mi (RWTH Aachen University Hospital - Medical Informatics)**

Mark O. Güld y Thomas M. Deserno [29] proponen un sistema de clasificación de imágenes médicas empleando el método del vecino más cercano (K-NN) con la ayuda de un umbral que decide entre dos posibles códigos que pueden darse.

- **Grupo MIRACLE**

Sara Lana Serrano, Julio Villena Román, Juan Carlos González Cristóbal y José Miguel Goñi Menoyo [35] proponen un sistema de clasificación de imágenes médicas resolviéndolo como un problema de aprendizaje máquina, sin tener en cuenta el dominio, ya que sus áreas de especialización no incluyen el tratamiento de imagen. Llevaron a cabo tres tipos de clasificadores que utilizan la imagen para predecir el código IRMA. Para llevar esto a cabo, todas las imágenes de entrenamiento, el desarrollo y el conjunto de datos fueron procesados mediante FIRE, que es un sistema de libre distribución de recuperación de imágenes basado en el contenido de las mismas. Proporciona además, todo un repertorio de características y de funciones de distancia. Específicamente, el paquete que se distribuye incluye un juego de scripts que extraen diferentes texturas de la imagen, incluyendo histogramas en color o niveles de gris, rasgos de Gabor o Tamura, descriptor de textura global, etc.

- **Grupo UNIBAS-DBIS**

Michael Sprigmann y Heiko Schuldt [56] proponen un sistema de clasificación de imágenes médicas basado en el Modelo de Distorsión de Imagen y un clasificador k-NN.

- **Grupo rwth_mi (RWTH Aachen University Hospital - Medical Informatics)**

Mark O. Güld, Christian Thies, Benedikt Fischer y Thomas M. Lehmann [30] proponen un sistema de clasificación de imágenes médicas basado en la extracción de características globales de la imagen y un clasificador k-NN. En ImageCLEF 2005, donde el objetivo era determinar a cuál de las posibles 57 clases pertenecía una imagen sin categorizar, el procedimiento de extracción de características fue similar, a excepción de considerar las características de Castelli y la distancia de Mahalanobis, y el clasificador empleado fue idéntico.

- **Grupo mtholyoke (Mount Holyoke Group)**

Desislava Petkova y Lisa Ballesteros [45] proponen un sistema de clasificación de imágenes médicas basado en la extracción de características globales de la imagen y un clasificador k-NN, aunque también analizan las prestaciones del algoritmo de las k-medias.

- **Grupo OHSU (Oregon Health and Science University)**

William Hersh, Jayashree Kalpathy-Cramer y Jeffery Jensen [31] proponen un sistema de clasificación de imágenes médicas empleando una arquitectura de perceptrón multicapa extrayendo previamente algunas características de las imágenes médicas.

- **Grupo OHSU (Oregon Health and Science University)**

Jayashree Kalpathy-Cramer y William Hersh [32] proponen un sistema de clasificación de imágenes médicas empleando una aproximación de un aprendizaje máquina supervisado utilizando principalmente características globales y una arquitectura de red neuronal.

- **Grupo BLOOM**

Tatiana Tommasi, Francesco Orabona y Barbara Caputo [59] proponen un sistema de clasificación de imágenes médicas buscando una representación de las mismas mediante descriptores locales y globales, que se combinan en la clasificación, la cual se basa en estrategias de SVM, y en la que presentan dos posibles alternativas. La estrategia que proponen es extraer un juego de características de cada imagen y emplear entonces, una Máquina de Vector Soporte (SVM) para clasificar las imágenes. Realizan un acercamiento local, utilizando descriptores SIFT y un acercamiento global, utilizando los píxeles en sí.

- **Grupo MU**

Bo Qiu, Chansheng Xu y Qi Tian [49] proponen un sistema de clasificación de imágenes médicas en la que se realizan dos pasos de clasificación independientes, empleando en cada uno de ellos SVM. En la primera fase se realiza una clasificación más general, y en la segunda, se busca un refinamiento para obtener una clasificación más exacta.

- **Grupo MedIC_CISMeF**

Florea, Rogozan, Cornea, Bensrhair y Darmoni [20] proponen un sistema de clasificación de imágenes médicas en la que se realiza una extracción de las características de las imágenes médicas una vez que la imagen se redimensiona, para utilizarlas en la implementación de un SVM.

- **Grupo MSRA**

Jing Liu, Yang Hu, Mingjing Li y Wei-Ying [39] Ma proponen un sistema de clasificación de imágenes médicas en la que se realiza una extracción de las características tanto globales como locales de las imágenes médicas, para utilizarlas en la implementación de dos SVM.

- **Grupo UFR**

Lokesh Setia, Alexandra Teinor, Alaa Halawani y Hans Burkhardt [55] proponen un sistema de clasificación de imágenes médicas en la que se realiza una extracción de las características locales correlativas de las imágenes médicas, para utilizarlas en la implementación de un SVM.

- **Grupo nctu_dblab**

Pei-Cheng Cheng, Been-Chian Chien, Hao-Ren Ke y Wei-Pang Yang [12] proponen un sistema de clasificación de imágenes médicas en la que se realiza una extracción de las características de las imágenes médicas, para utilizarlas en la implementación de un SVM.

- **Grupo cind**

Md. Mahmudur Rahman, Bipin C. Desai, y Prabir Bhattacharya [51] proponen un sistema de clasificación de imágenes médicas en la que se realiza una extracción de características globales de las imágenes médicas, para utilizarlas en la implementación de un SVM.

- **Grupo infocomm (Institute of Infocomm)**

Bo Qiu, Wei Xiong Qi Tian y Chang Sheng Xu [50] proponen un sistema de clasificación de imágenes médicas en la que se realiza una extracción de características tanto globales de bajo nivel como de regiones de las imágenes médicas, para utilizarlas en la implementación de un SVM.

3.- ANÁLISIS DEL SISTEMA DE CLASIFICACIÓN DE IMÁGENES

3.1.- Introducción

Tal y como se ha comentado en el capítulo 2, el diseño de un sistema de clasificación de imágenes basado en contenido depende de la aplicación concreta para la que se vaya a utilizar. Durante el diseño del mismo se toman una serie de decisiones:

1. Conjunto de características elegidas para identificar la imagen (una vez se han probado las características se decide cuál de ellas es más representativa para el tipo de imágenes con el que se va a trabajar). El conjunto de características elegido se va a explicar en el presente capítulo.
2. Algoritmo de clasificación de imágenes. El algoritmo de clasificación elegido se va a expresar en el presente capítulo.
3. Estructura de la codificación (indexación) de las imágenes. La estructura de codificación que se va a utilizar, tal y como se explicó en el capítulo del estado del arte, es la codificación IRMA.
4. Estructura de la base de datos utilizada. La base de datos que se va a utilizar se detalló en el capítulo del estado del arte.
5. Conjunto de imágenes de imágenes de entrenamiento (previamente clasificadas) y de test. El conjunto de imágenes utilizado consta de dos conjuntos, uno de test y otro de entrenamiento, tal y como se detalló en el capítulo del estado del arte.

La base de datos que se va a usar en este proyecto ha sido proporcionada por CLEF (*Cross Language Evaluation Forum*), y está formada por un conjunto de imágenes médicas en formato .PNG perteneciente al grupo IRMA de la *Aachen University of Technology* en Alemania y son empleadas en el proyecto IRMA, cuyo objetivo, como se ha visto anteriormente, es el desarrollo y aplicación de métodos de alto nivel para la recuperación de imágenes médicas basado en el contenido de las mismas, como ayuda al diagnóstico de pacientes. El conjunto de imágenes con el que se va a trabajar es un conjunto de 10.000 imágenes de entrenamiento previamente clasificadas y 1.000 imágenes de test cuya clasificación se desconoce y deben ser etiquetadas por el sistema implementado.

El sistema que se va a diseñar constará de dos bloques fundamentales, uno de extracción de características de las imágenes, y otro de búsqueda de imágenes con características similares a la que se está analizando, en el que se compararán las

características de las imágenes previamente clasificadas y de las que se desconoce la clase a la que pertenecen para así etiquetarlas.

3.2.- Extracción de características

Para realizar el estudio de las diferentes características y obtener las más significantes para la aplicación se han seleccionado cinco imágenes de las proporcionadas por el conjunto de entrenamiento de la base de datos. Estas imágenes son las mostradas en la siguiente figura.

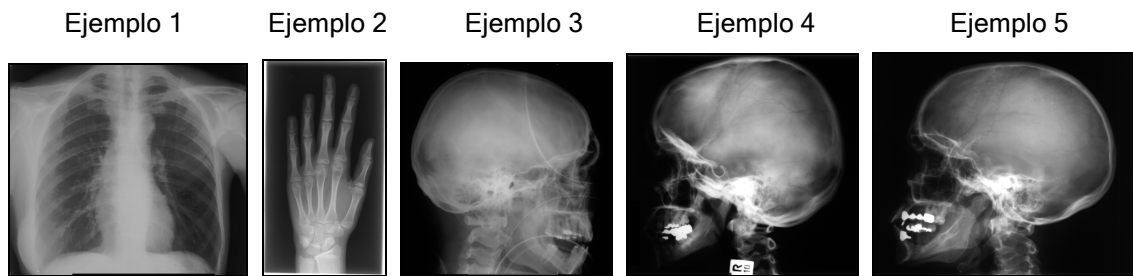


Figura 13. Imágenes de ejemplo

La selección de las imágenes que se van a estudiar como ejemplos para ver las características más significativas se han elegido de manera aleatoria, con la única excepción del 'Ejemplo4.png' que se eligió para que fuera fácilmente confundible con la imagen 'Ejemplo3.png', de manera que se pudieran identificar las características y parámetros de las mismas que permitieran diferenciarlas, y el 'Ejemplo5.png' que se eligió porque su código debería coincidir con el código de la imagen 'Ejemplo4.png', de manera que se pueden estudiar las características y parámetros que permiten reconocer imágenes con el mismo código.

Para realizar el análisis de las imágenes anteriores se realiza un estudio de algunas de las características explicadas en el apartado del estado del arte y los parámetros que pueden ser variados en cada una de ellas, que constituirán la información que se tendrá disponible para diferenciar las imágenes entre sí y proceder a la posterior creación de la base de datos con el conjunto de imágenes de entrenamiento, y la clasificación del conjunto de imágenes de test.

Para el estudio de las características de las imágenes de forma global, todas ellas se redimensionan al mismo tamaño mediante interpolación lineal. El tamaño elegido para esta redimensión ha sido 256x256 píxeles. Posteriormente, para su estudio local, se divide cada imagen en un número arbitrario de secciones de igual tamaño.

1.- HISTOGRAMA

Se va a analizar el histograma de las imágenes de ejemplo con 128 niveles de gris diferentes. Al ser las imágenes en blanco y negro se ha considerado suficiente como número de

3.- Análisis del sistema de clasificación de imágenes

niveles de la escala de grises para representar cada imagen. Los resultados son los mostrados a continuación.

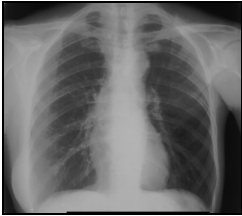
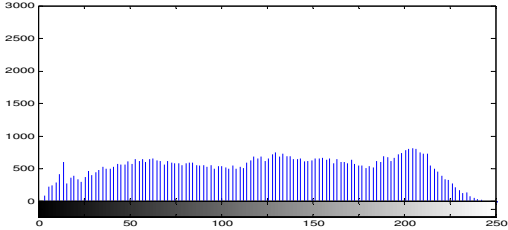

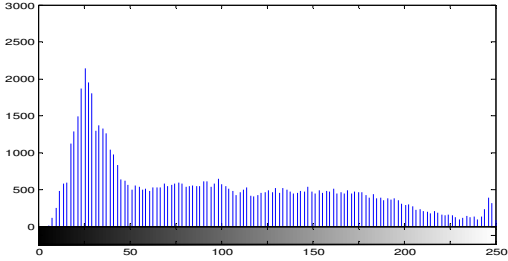
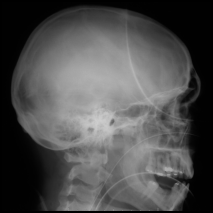
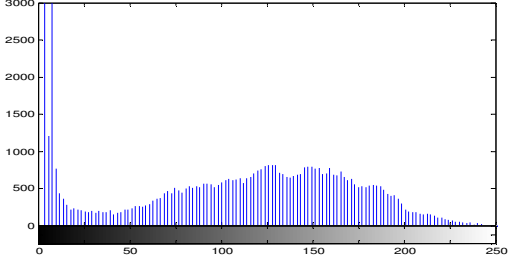

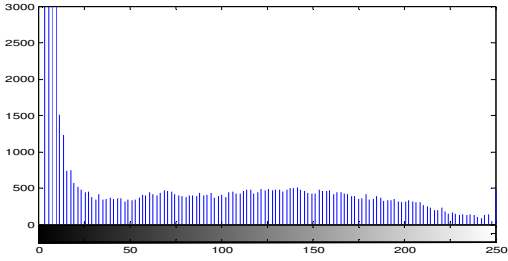
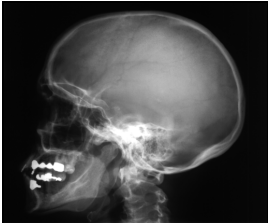
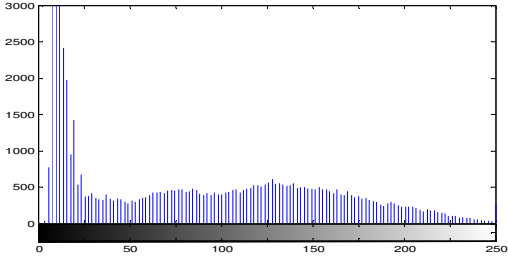
Imagen de ejemplo	Histograma
	
	
	
	
	

Tabla 3. Histograma de imágenes de ejemplo

3.- Análisis del sistema de clasificación de imágenes

Además de los resultados gráficos, también se han obtenido resultados numéricos, ya que se almacena el valor concreto de la frecuencia de cada uno de los niveles de gris (estos valores serán los que permitirán comparar cuantitativamente unas imágenes con otras, ya que la representación gráfica sólo nos permite realizar una comparación cualitativa).

Se observa que para las tres primeras imágenes los valores de las frecuencias de los distintos valores de grises se diferencian considerablemente, mientras que en los tres últimos casos el histograma es bastante similar, ya que la imagen es la misma, aunque en el caso del 'Ejemplo3.png' tiene otra orientación.

De los resultados obtenidos se puede deducir que esta característica puede ser útil para diferenciar imágenes de diferentes partes del cuerpo, pero no imágenes de la misma parte del cuerpo con distinta orientación.

El parámetro que puede ser variado en este caso es el número de niveles de gris a tener en cuenta. Evidentemente cuanto mayor sea este número más precisión se tendrá a la hora de diferenciar los niveles de gris y más fácil será diferenciar unas imágenes de otras.

2.- ESTADÍSTICOS

Se van a analizar los estadísticos de las imágenes de ejemplo. Los resultados son los mostrados a continuación.



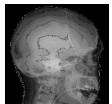
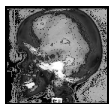
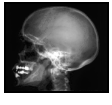
Imagen de Ejemplo	Media	Mediana	Varianza	Asimetría	Curtosis
	125,725	129	3943,2	-0,093	1,863
	97,348	85	4394,741	0,53	2,12
	97,652	108	4592,029	-0,082	1,718
	86,273	73	5562,282	0,473	1,922
	81,534	69	4842,362	0,529	2,006

Tabla 4. Estadísticos de imágenes de ejemplo

3.- Análisis del sistema de clasificación de imágenes

Observando los resultados obtenidos en la tabla anterior, se puede determinar que vistas una a una por separado, cada una de las características no aporta demasiada información para diferenciar las imágenes, ya que, por ejemplo, la media de las imágenes 'Ejemplo2.png' y 'Ejemplo3.png' es bastante parecida, aunque estas dos imágenes no se corresponderían con el mismo código. Sin embargo, si se observa el conjunto de las cinco imágenes, se ve que para las dos últimas las características son bastante más parecidas que en el caso de las tres primeras.

En la tabla anterior se puede observar cómo los valores de los estadísticos de primer orden obtenidos tienen valores muy distintos para las tres primeras imágenes, mientras que son bastante parecidos en el caso de las dos últimas imágenes que estarían asociadas al mismo código.

En esta característica no existe ningún parámetro que pueda variarse para mejorar el diseño del sistema de clasificación.

3.- DIMENSIÓN FRACTAL

Se va a analizar la dimensión fractal de las imágenes de ejemplo. Los resultados son los mostrados a continuación.

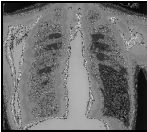

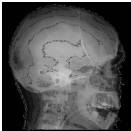
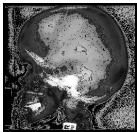
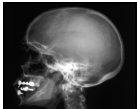
Imagen de Ejemplo	Dimensión Fractal
	0,9363
	0,9131
	0,9136
	0,9023
	0,8969

Tabla 5. Dimensión fractal de imágenes de ejemplo

3.- Análisis del sistema de clasificación de imágenes

La variación del valor de esta característica entre imágenes distintas entre sí, e imágenes que deben tener la misma codificación no es demasiado grande, ya que la varianza de este valor es muy pequeña y toma valores muy cercanos entre sí para cualquier imagen.

En este caso tampoco existe ningún parámetro que pueda variarse para obtener mejores resultados.

4.- CARACTERÍSTICAS DE GABOR

Se van a analizar las características de Gabor de las imágenes de ejemplo. Las máscaras de los filtros usados, con distintas escalas y orientación, se representan en la siguiente tabla.


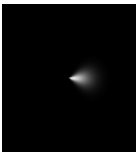
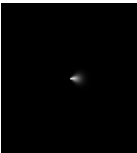
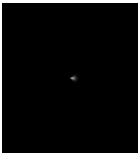


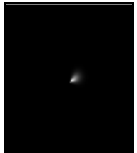


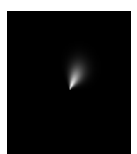




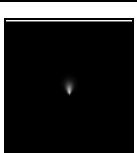
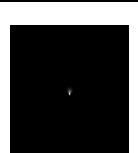
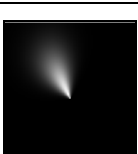

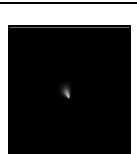
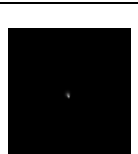

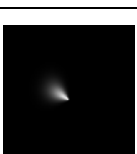
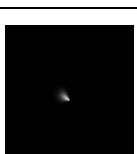
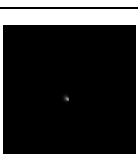
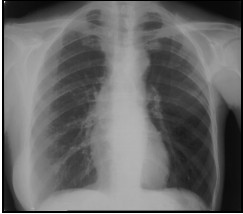

		ESCALA			
		1	2	3	4
ORIENTACIÓN	1				
	2				
	3				
	4				
	5				
	6				

Tabla 6. Filtros de Gabor para 24 combinaciones de escalas y orientaciones


3.- Análisis del sistema de clasificación de imágenes


La elección de seis orientaciones y cuatro escalas se ha hecho de manera que se llegara a un compromiso entre unos resultados aceptables y un coste computacional no demasiado elevado, ya que este último aumenta proporcionalmente al número de filtros por el que vayan a pasar las imágenes. De esta manera tendremos 24 máscaras de filtros distintas, a través de las cuales se filtran las imágenes. Posteriormente se calculan dos estadísticos, la media y la varianza, de las imágenes filtradas, de manera que hay 48 valores que representan esta característica para cada una de ellas, por lo que el coste computacional de esta característica es grande comparado, por ejemplo, con el de la dimensión fractal.

Se han obtenido la media y la varianza de las 24 imágenes filtradas de cada una de las imágenes originales con las que se está trabando. Los valores obtenidos se muestran en la siguiente tabla.

Imagen de Ejemplo		Características					
	Orientación			Escala			
				1	2	3	4
		1	Media	0.1107	0.0754	0.0471	0.0432
			Varianza	0.0094	0.0097	0.0076	0.0117
		2	Media	0.0436	0.0573	0.1096	0.0746
			Varianza	0.0068	0.0042	0.0089	0.0096
		3	Media	0.0461	0.0416	0.0424	0.0563
			Varianza	0.0076	0.0117	0.0067	0.0039
		4	Media	0.1062	0.0725	0.0439	0.0378
			Varianza	0.0071	0.0094	0.0070	0.0102
		5	Media	0.0395	0.0540	0.0958	0.0668
			Varianza	0.0067	0.0036	0.0057	0.0081
		6	Media	0.0398	0.0302	0.0339	0.0483
			Varianza	0.0057	0.0063	0.0065	0.0037
	Orientación			Escala			
				1	2	3	4
		1	Media	0.1228	0.0517	0.0316	0.0598
			Varianza	0.0136	0.0073	0.0066	0.0072
		2	Media	0.0441	0.0712	0.1210	0.0508
			Varianza	0.0073	0.0144	0.0128	0.0069
		3	Media	0.0309	0.0587	0.0433	0.0699
			Varianza	0.0065	0.0069	0.0074	0.0136
		4	Media	0.1147	0.0478	0.0293	0.0561
			Varianza	0.0117	0.0064	0.0063	0.0065
		5	Media	0.0414	0.0655	0.0978	0.0405
			Varianza	0.0075	0.0117	0.0096	0.0062
		6	Media	0.0259	0.0508	0.0372	0.0532
			Varianza	0.0056	0.0057	0.0070	0.0079

3.- Análisis del sistema de clasificación de imágenes

				Escala			
				1	2	3	4
	Orientación	1	Media	0.1348	0.0560	0.0491	0.0816
			Varianza	0.0078	0.0088	0.0090	0.0134
		2	Media	0.0405	0.0524	0.1343	0.0551
			Varianza	0.0068	0.0076	0.0072	0.0085
		3	Media	0.0479	0.0803	0.0391	0.0513
			Varianza	0.0084	0.0119	0.0060	0.0072
		4	Media	0.1332	0.0533	0.0449	0.0771
			Varianza	0.0062	0.0079	0.0074	0.0358
		5	Media	0.0493	0.1305	0.0501	0.0092
			Varianza	0.0047	0.0061	0.0049	0.0394
6	Media	0.0715	0.0305	0.0460	0.0072		
	Varianza	0.0062	0.0066	0.0029	0.0037		

				Escala			
				1	2	3	4
	Orientación	1	Media	0.1388	0.0625	0.0537	0.0880
			Varianza	0.0135	0.0058	0.0121	0.0111
		2	Media	0.0595	0.0644	0.1382	0.0613
			Varianza	0.0114	0.0066	0.0128	0.0053
		3	Media	0.0518	0.0848	0.0578	0.0634
			Varianza	0.0110	0.0110	0.0112	0.0058
		4	Media	0.1369	0.0586	0.0478	0.0798
			Varianza	0.0115	0.0046	0.0083	0.0089
		5	Media	0.0541	0.0611	0.1338	0.0531
			Varianza	0.0103	0.0045	0.0090	0.0041
		6	Media	0.0406	0.0734	0.0467	0.0563
			Varianza	0.0042	0.0072	0.0086	0.0037


				Escala			
				1	2	3	4
	Orientación	1	Media	0.1271	0.0608	0.0501	0.0782
			Varianza	0.0121	0.0093	0.0100	0.0117
		2	Media	0.0612	0.0623	0.1265	0.0596
			Varianza	0.0097	0.0067	0.0112	0.0085
		3	Media	0.0484	0.0769	0.0598	0.0611
			Varianza	0.0087	0.0100	0.0087	0.0064
		4	Media	0.1251	0.0572	0.0450	0.0739
			Varianza	0.0096	0.0071	0.0066	0.0071
		5	Media	0.0563	0.0586	0.1222	0.0530
			Varianza	0.0069	0.0057	0.0064	0.0061
		6	Media	0.0395	0.0684	0.0501	0.0548
			Varianza	0.0042	0.0054	0.0052	0.0053

Tabla 7. Características de Gabor de imágenes de ejemplo

3.- Análisis del sistema de clasificación de imágenes

Con los resultados anteriores se puede llegar a la conclusión de que, a medida que aumenta la escala del filtro aplicado, el valor de la media y la varianza disminuyen. También vemos que para las orientaciones horizontales, los valores de los estadísticos son mayores.

Los parámetros que en este caso se pueden variar son el número de escalas y orientaciones para los que se calculan las diferentes máscaras de los filtros por los que posteriormente pasarán las imágenes.

5.- CARACTERÍSTICAS DE LAWS

Se han obtenido imágenes filtradas a partir de las imágenes de ejemplo originales con las diferentes máscaras que se han visto previamente en el capítulo del estado del arte (14 texturas diferentes). Además de las imágenes filtradas se ha obtenido la energía de cada una de ellas. Los valores obtenidos se muestran en la siguiente tabla:

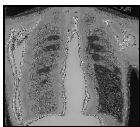







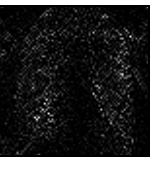
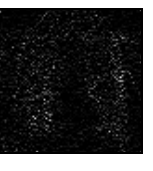
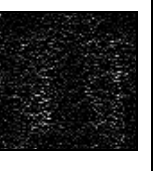










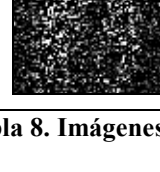


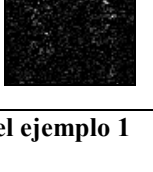

Imagen Ejemplo	Imágenes filtradas					
		L5	E5	S5	W5	R5
	L5					
	E5					
	S5					
	W5					
	R5					

Tabla 8. Imágenes filtradas con máscaras de Laws del ejemplo 1







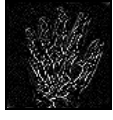
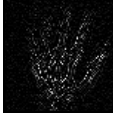
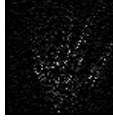
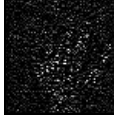

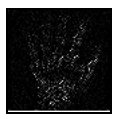
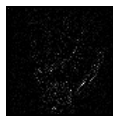
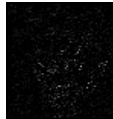


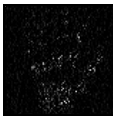



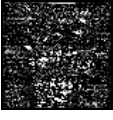


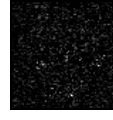

Imagen Ejemplo	Imágenes filtradas					
		L5	E5	S5	W5	R5
	L5					
	E5					
	S5					
	W5					
	R5					

Tabla 9. Imágenes filtradas con máscaras de Laws del ejemplo 2

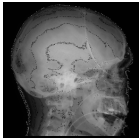





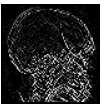
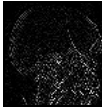
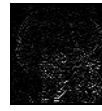



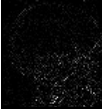

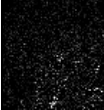


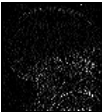
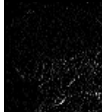






Imagen Ejemplo	Imágenes filtradas					
		L5	E5	S5	W5	R5
	L5					
	E5					
	S5					
	W5					
	R5					

Tabla 10. Imágenes filtradas con máscaras de Laws de ejemplo 3

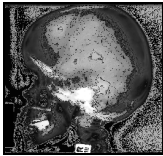



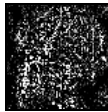











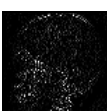








Imagen Ejemplo	Imágenes filtradas					
		L5	E5	S5	W5	R5
	L5					
	E5					
	S5					
	W5					
	R5					

Tabla 11. Imágenes filtradas con filtros de Laws del ejemplo 4





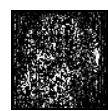



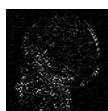


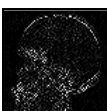

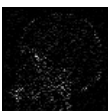







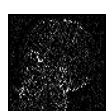


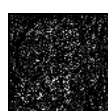
Imagen Ejemplo	Imágenes filtradas					
		L5	E5	S5	W5	R5
	L5					
	E5					
	S5					
	W5					
	R5					

Tabla 12. Imágenes filtradas con filtros Laws del ejemplo 5

3.- Análisis del sistema de clasificación de imágenes

En las tablas anteriores se han presentado las imágenes filtradas con las máscaras de Laws de las imágenes de ejemplo.

Una vez hemos visto las imágenes resultado de filtrar las imágenes de ejemplo originales con las máscaras de Laws, se combinan las características que son similares entre sí (máscaras sensibles a los bordes verticales y horizontales pueden unirse en una sola característica) tal y como se ha visto anteriormente. Una vez las características están combinadas se obtiene la energía de las imágenes, que se presenta en la siguiente tabla.

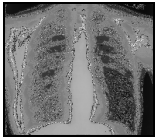

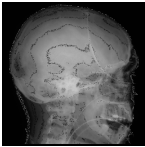
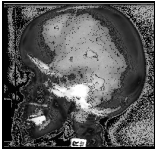
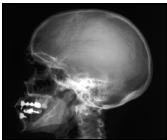
Imagen Ejemplo	Energía					
		L5	E5	S5	W5	R5
	L5					
	E5	144.2040	2.6439			
	S5	13.7771	0.5561	0.1608		
	W5	8.3940	0.3438	0.1165	0.1009	
	R5	25.2248	0.7602	0.3044	0.3269	1.4439
		L5	E5	S5	W5	R5
	L5					
	E5	381.5290	4.5645			
	S5	57.6254	0.8653	0.2156		
	W5	42.2901	0.5423	0.1644	0.1539	
	R5	170.7232	1.4239	0.5399	0.6260	3.1378
		L5	E5	S5	W5	R5
	L5					
	E5	270.6000	7.0123			
	S5	27.6636	1.3455	0.3657		
	W5	12.9248	0.8062	0.2719	0.2426	
	R5	26.5489	1.8419	0.7370	0.7998	3.3751
		L5	E5	S5	W5	R5
	L5					
	E5	518.0757	12.1672			
	S5	68.8376	2.1771	0.5277		
	W5	38.6261	1.1791	0.3479	0.2797	
	R5	144.7763	2.2452	0.8113	0.8445	3.5692
		L5	E5	S5	W5	R5
	L5					
	E5	536.7944	15.4214			
	S5	55.2920	2.7196	0.6737		
	W5	23.7442	1.4647	0.4564	0.3825	
	R5	43.3458	2.9214	1.1233	1.1682	4.6680
		L5	E5	S5	W5	R5

Tabla 13. Energía de las imágenes filtradas con matrices de características similares

3.- Análisis del sistema de clasificación de imágenes

Los valores de energía obtenidos para imágenes pertenecientes a la misma clase son muy parecidos, mientras que son totalmente diferentes a los obtenidos para imágenes distintas, quedando en un punto intermedio la imagen 'Ejemplo3.png', que es muy similar a las imágenes 'Ejemplo4.png' y 'Ejemplo5.png', con la salvedad de que su orientación no es la misma.

El grado de decisión dentro de esta característica es la agrupación de las diferentes máscaras con que se filtrarán las imágenes para obtener la energía.

6.- CARACTERÍSTICAS DE TAMURA

Se han extraído tres de las características de Tamura que se han explicado en el capítulo del estado del arte, la tosquedad, el contraste y la direccionalidad. Además de las imágenes filtradas se ha obtenido la energía de cada una de ellas (y del histograma en el caso de la direccionalidad). Los valores obtenidos se muestran en la siguiente tabla.

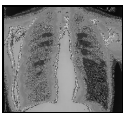
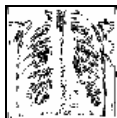




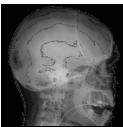


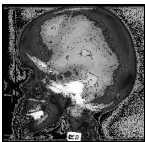


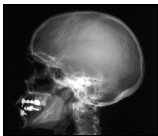


Imagen de Ejemplo	Tosquedad	Contraste	Direccionalidad
	6.6678	0.2115	3.2878
			
	6.7888	0.2155	5.8989
			
	6.2001	0.2327	3.8953
			
	6.2040	0.2485	4.0971
			
	5.9744	0.2289	5.2673
			

Tabla 14. Características de Tamura de las imágenes de ejemplo

3.- Análisis del sistema de clasificación de imágenes

Vemos cómo el valor de la tosquedad no da demasiada información ya que su valor es muy parecido en todos los casos; sin embargo, los valores del contraste y de la direccionalidad, son similares en el caso de las imágenes que pertenecen a la misma clase, de manera que estás dos características pueden ayudar mucho a la hora de decidir.

No existe ningún parámetro que se pueda variar en el estudio de esta característica.

7.- MATRIZ DE CO-OCURRENCIA

Se han extraído los estadísticos obtenidos tras el cálculo de la matriz de co-ocurrencia de cada una de las imágenes de ejemplo con las que se está trabajando. Los valores obtenidos se muestran en la siguiente tabla.



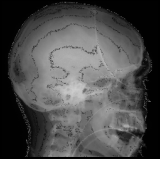
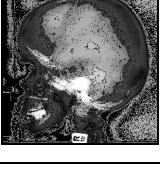

Imagen Ejemplo	Orient.	Energía	Contraste	Entropía	Homogeneidad	Correlación
	0º	1.7253	0.2308	0.0872	0.9566	0.9570
	45º	1.8070	0.2191	0.1221	0.9422	0.9394
	90º	1.6922	0.2370	0.0812	0.9625	0.9597
	135º	1.8046	0.2195	0.1214	0.9426	0.9398
	0º	1.8824	0.1968	0.1165	0.9421	0.9534
	45º	1.9315	0.1907	0.1409	0.9305	0.9430
	90º	1.7911	0.2109	0.0835	0.9588	0.9667
	135º	1.9640	0.1855	0.1556	0.9235	0.9370
	0º	1.6913	0.2264	0.0775	0.9615	0.9696
	45º	1.7876	0.2128	0.1187	0.9446	0.9534
	90º	1.7077	0.2249	0.0878	0.9593	0.9656
	135º	1.7832	0.2136	0.1174	0.9455	0.9539
	0º	1.7658	0.2262	0.0843	0.9584	0.9727
	45º	1.9027	0.2118	0.1596	0.9332	0.9481
	90º	1.8294	0.2218	0.1262	0.9472	0.9590
	135º	1.9029	0.2109	0.1572	0.9331	0.9489
	0º	1.7233	0.2414	0.0904	0.9553	0.9679
	45º	1.8043	0.2315	0.1276	0.9399	0.9547
	90º	1.7190	0.2427	0.0903	0.9559	0.9679
	135º	1.8116	0.2298	0.1291	0.9384	0.9541

Tabla 15. Estadísticos de la matriz de co-ocurrencia de las imágenes de ejemplo

Se han elegido cuatro valores típicos de orientación, de manera que los resultados obtenidos sean aceptables mientras que el coste computacional del cálculo de esta característica en cada imagen, que crece proporcionalmente con el número de direcciones analizadas, no sea demasiado elevado.


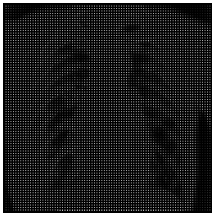
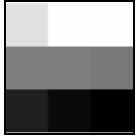

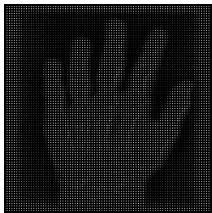


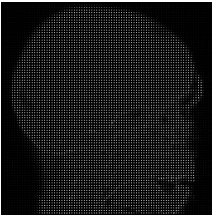

3.- Análisis del sistema de clasificación de imágenes

El valor de los estadísticos obtenidos a partir de la matriz de co-ocurrencia, que se muestran en la tabla anterior, conduce a las siguientes conclusiones: el valor de la homogeneidad y la correlación es bastante similar en todas las imágenes, de manera que de forma individual no proporcionarían demasiada información, aunque completan al resto de parámetros, siendo entre todos ellos el contraste el que proporciona mayor información a la hora de decidir.

El parámetro que se puede variar en esta característica para optimizar el funcionamiento del sistema es el número de orientaciones que se estudian para cada imagen. Evidentemente cuando mayor sea el número de orientaciones estudiadas mejor será el resultado, pero el coste computacional del estudio de cada una de estas orientaciones es muy elevado, y es por esa razón por la que se ha decidido estudiar cuatro orientaciones, que da información suficiente como para ayudar a la tarea de decisión, pero con un coste computacional aceptable.

8.- TRANSFORMADA DISCRETA DEL COSENO

Se han extraído los coeficientes de la transformada discreta del coseno de cada una de las imágenes de ejemplo con las que se está trabajando. Los valores obtenidos y las imágenes transformadas se muestran en la siguiente tabla.

Imagen de Ejemplo	Imagen Transformada	Componentes baja frecuencia									
		 <table><tr><td>225</td><td>255</td><td>254</td></tr><tr><td>127</td><td>124</td><td>121</td></tr><tr><td>31</td><td>10</td><td>0</td></tr></table>	225	255	254	127	124	121	31	10	0
225	255	254									
127	124	121									
31	10	0									
		 <table><tr><td>255</td><td>251</td><td>243</td></tr><tr><td>6</td><td>9</td><td>16</td></tr><tr><td>2</td><td>0</td><td>0</td></tr></table>	255	251	243	6	9	16	2	0	0
255	251	243									
6	9	16									
2	0	0									
		 <table><tr><td>242</td><td>248</td><td>254</td></tr><tr><td>4</td><td>0</td><td>2</td></tr><tr><td>2</td><td>1</td><td>0</td></tr></table>	242	248	254	4	0	2	2	1	0
242	248	254									
4	0	2									
2	1	0									


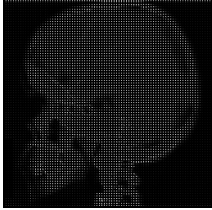

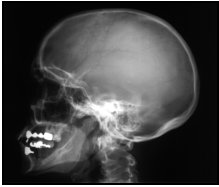
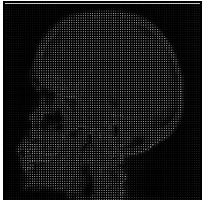
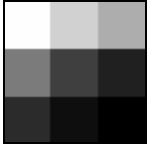
		 <table> <tr><td>254</td><td>244</td><td>238</td></tr> <tr><td>46</td><td>44</td><td>46</td></tr> <tr><td>0</td><td>0</td><td>5</td></tr> </table>	254	244	238	46	44	46	0	0	5
254	244	238									
46	44	46									
0	0	5									
		 <table> <tr><td>255</td><td>209</td><td>173</td></tr> <tr><td>123</td><td>63</td><td>32</td></tr> <tr><td>43</td><td>14</td><td>0</td></tr> </table>	255	209	173	123	63	32	43	14	0
255	209	173									
123	63	32									
43	14	0									

Tabla 16. Coeficientes de la transformada discreta del coseno de las imágenes de ejemplo

Se puede observar claramente cómo las componentes de baja frecuencia de las imágenes que pertenecen a la misma parte del cuerpo (en este caso el cráneo) son muy parecidas, aunque existe un margen de diferencia suficiente como para permitirnos diferenciar las diferentes orientaciones en distintas imágenes.

9.- TRANSFORMADA DE WAVELET DISCRETA

Se han extraído los estadísticos (media y desviación típica) de la transformada discreta de wavelet en las bandas HL, LH y HH de cada una de las imágenes de ejemplo con las que se está trabajando. Los valores obtenidos se muestran en la siguiente tabla.






Imagen Ejemplo	Banda HL		Banda LH		Banda HH	
	Media		Media		Media	
	Media	0.000026	Media	-0.004461	Media	0.00057
	Desv. típica	0.238618	Desv. típica	0.090166	Desv. típica	0.043888
	Media	-0.001792	Media	0.004246	Media	-0.001258
	Desv. típica	0.095961	Desv. típica	0.107217	Desv. típica	0.040260
	Media	-0.000855	Media	0.000433	Media	-0.000381
	Desv. típica	0.206070	Desv. típica	0.091919	Desv. típica	0.061855
	Media	-0.014068	Media	-0.003180	Media	-0.000747
	Desv. típica	0.127576	Desv. típica	0.106615	Desv. típica	0.073888
	Media	0.004581	Media	0.004410	Media	0.000573
	Desv. típica	0.164287	Desv. típica	0.113360	Desv. típica	0.082368

Tabla 17. Estadísticos de la transformada discreta de Wavelet de las imágenes de ejemplo

3.- Análisis del sistema de clasificación de imágenes

En vista de los resultados mostrados en la tabla anterior, se observa que esta característica no aporta demasiada información a la hora de decidir ya que los valores de la media y la varianza en las distintas bandas para las diferentes imágenes son bastante parecidos.

En esta característica no hay ningún parámetro que se pueda variar para optimizar sus resultados.

10.- TRANSFORMADA DISCRETA DE FOURIER

Se han extraído los módulos de los coeficientes de la transformada discreta de Fourier de cada una de las imágenes de ejemplo con las que se está trabajando. Los valores obtenidos y las imágenes transformadas se muestran en la siguiente tabla.


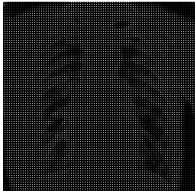


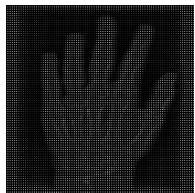


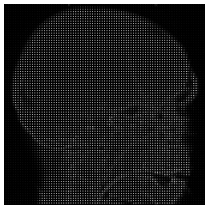


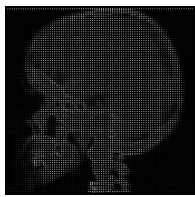


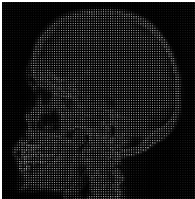
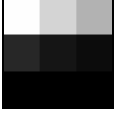
Imagen de Ejemplo	Imagen Transformada	Componentes baja frecuencia									
		 <table> <tr> <td>217</td><td>255</td><td>254</td></tr> <tr> <td>79.48</td><td>84.0536</td><td>88.73</td></tr> <tr> <td>15.26</td><td>10</td><td>0</td></tr> </table>	217	255	254	79.48	84.0536	88.73	15.26	10	0
217	255	254									
79.48	84.0536	88.73									
15.26	10	0									
		 <table> <tr> <td>255</td><td>250</td><td>243</td></tr> <tr> <td>1.41</td><td>3.6</td><td>8.54</td></tr> <tr> <td>0</td><td>1</td><td>2.24</td></tr> </table>	255	250	243	1.41	3.6	8.54	0	1	2.24
255	250	243									
1.41	3.6	8.54									
0	1	2.24									
		 <table> <tr> <td>241</td><td>248</td><td>255</td></tr> <tr> <td>0</td><td>4.47</td><td>2</td></tr> <tr> <td>17</td><td>19.31</td><td>17</td></tr> </table>	241	248	255	0	4.47	2	17	19.31	17
241	248	255									
0	4.47	2									
17	19.31	17									
		 <table> <tr> <td>255</td><td>245</td><td>239</td></tr> <tr> <td>2</td><td>0</td><td>3.61</td></tr> <tr> <td>59.48</td><td>56.04</td><td>52.39</td></tr> </table>	255	245	239	2	0	3.61	59.48	56.04	52.39
255	245	239									
2	0	3.61									
59.48	56.04	52.39									
		 <table> <tr> <td>255</td><td>212</td><td>178.03</td></tr> <tr> <td>78.39</td><td>39.96</td><td>20.81</td></tr> <tr> <td>25.55</td><td>9.22</td><td>0</td></tr> </table>	255	212	178.03	78.39	39.96	20.81	25.55	9.22	0
255	212	178.03									
78.39	39.96	20.81									
25.55	9.22	0									

Tabla 18. Módulo de los coeficientes de la transformada discreta de Fourier de las imágenes de ejemplo

3.- Análisis del sistema de clasificación de imágenes

En vista de los resultados mostrados en la tabla anterior se puede afirmar que el módulo de los coeficientes de la transformada discreta de Fourier no nos ayuda demasiado a la hora de decidir, ya que su valor no es claramente diferente para imágenes que pertenecen a clases diferentes.

11.- RESUMEN

Cada imagen se analiza de forma independiente a las demás, y sus características (globales, locales y de búsqueda de objetos) se almacenan en un vector, que posteriormente se usará para comparar las imágenes entre sí y realizar la clasificación. Las características almacenadas en este vector son las siguientes (suponiendo que se ha dividido cada imagen redimensionada a 256x256 píxeles en 4 trozos de 128x128 píxeles cada uno para obtener las características locales):

Característica	Número Posiciones Característica Global	Número Posiciones Características Locales
Histograma	128	$4 * 128 = 512$
Estadísticos	5	$4 * 5 = 20$
Dimensión fractal	1	$4 * 1 = 4$
Filtros de Gabor	48	$4 * 48 = 192$
Filtros de Laws	14	$4 * 14 = 56$
Filtros de Tamura	3	$4 * 3 = 12$
Matriz de co-ocurrencia	20	$4 * 20 = 80$
Transformada Discreta del Coseno	9	$4 * 9 = 36$
Transformada Discreta de Wavelet	6	$4 * 6 = 24$
Transformada Discreta de Fourier	9	$4 * 9 = 36$
Total	243	$4 * 243 = 972$

Tabla 19. Características de cada imagen

Con el conjunto de imágenes de entrenamiento, sus características y el código IRMA que tiene cada una de ellas asociado se va a generar el modelo del sistema de clasificación de imágenes basado en contenido. A partir de este modelo el sistema debe ser capaz de clasificar las imágenes pertenecientes al conjunto de test mediante la comparación de las características de las mismas.

3.3.- Búsqueda de imágenes (Algoritmo k-NN)

Este es el bloque fundamental del sistema, puesto que es el que le dota de "inteligencia", ya que a partir de los conocimientos que se le proporcionan, en este caso el conjunto de imágenes de entrenamiento, sus características y el código asociado, va a ser capaz de clasificar otro conjunto de imágenes que no tienen código previamente asociado mediante un proceso de aprendizaje.

3.- Análisis del sistema de clasificación de imágenes

Al igual que en el caso de las imágenes de entrenamiento, se van a obtener las características de cada una de las imágenes que pertenecen al conjunto de entrenamiento. El procedimiento seguido es el mismo que en el caso del conjunto de imágenes de test, con la diferencia de que no tienen código previamente asociado, sino que a partir del algoritmo de clasificación es el sistema el que tiene que asociarles uno.

El algoritmo escogido para implementar el módulo de búsqueda de imágenes y su posterior clasificación es, tal y como se ha comentado en el capítulo del estado del arte, el algoritmo del “vecino más cercano” k-Nearest Neighbour, que es un algoritmo de aprendizaje supervisado y no paramétrico en el que, dado un conjunto de ejemplos en forma de pares entrada-salida (conjunto de imágenes de entrenamiento) es capaz de encontrar un conjunto de reglas para relacionar las entradas con las salidas.

La idea en la que se basa este algoritmo es que la probabilidad de que dos elementos cuyas características son parecidas pertenezcan a la misma clase es alta.

Para explicar el funcionamiento del algoritmo implementado se va a procesar una imagen del conjunto de test. Para ello previamente se ha debido generar la base de datos con el conjunto de imágenes de entrenamiento. En este proyecto se van a estudiar diferentes combinaciones de las características previamente estudiadas, pero en esta ocasión, para explicar el funcionamiento del algoritmo del vecino más cercano se va a usar el caso en el que se incluyen todas las características, cada una de ellas con el mismo peso que el resto.

La imagen de test escogida es la mostrada en la figura siguiente.



Código IRMA asociado: 1121-220-230-700

Figura 14. Imagen de ejemplo algoritmo k-NN

Una vez que se tiene la base de datos y la imagen de test que se va a procesar, se ha de elegir el número de vecinos con el que se quiere trabajar. Hay que elegir este número de manera que no debe ser muy pequeño, ya que si la vecindad es muy pequeña no contendrá ningún punto de los datos, debe ser lo suficientemente grande como para que la estimación realizada tenga significado y así la probabilidad de clasificación errónea sea más pequeña. Pero tampoco puede ser excesivamente grande, ya que en ese caso los puntos muestrales no estarían lo suficientemente cerca como para que la clasificación fuera fiable.

Se van a probar los resultados de la clasificación de la imagen de test escogida para distintos números de vecinos más cercanos. Como puede suceder que haya empate entre las

3.- Análisis del sistema de clasificación de imágenes

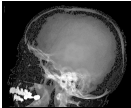
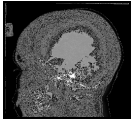
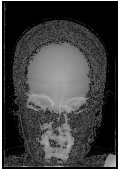
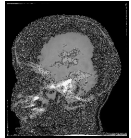
probabilidades de ocurrencia de dos o más códigos de los asociados a los k vecinos más cercanos, hay que definir una regla para decidir entre ellos. Además de la definición de esta regla se elige k impar para reducir la probabilidad de empate.

En caso de que varias clases tengan la misma probabilidad de ocurrencia, se asignará como correcta la que pertenezca al elemento que menor distancia presenta con el elemento que se está estudiando.

Si se observa el reparto de imágenes entre las diferentes clases (Apéndice A) se puede ver como hay clases que tienen muchos más elementos muestrales que otras, e incluso hay clases en las que se encuentra un número muy pequeño de individuos (por ejemplo 10). Por este motivo, el valor de k (número de vecinos más cercanos) más alto estudiado es $k = 7$, ya que valores mayores no tendrían sentido para estas clases con pocos individuos porque aunque la menor distancia fuera la que tiene con cada uno de los elementos de esa clase, el resto de los vecinos más cercanos es más que probable que no pertenecieran a la clase correcta porque no habría más elementos en ella.

La distancia escogida para la implementación del algoritmo, por su sencillez y los buenos resultados obtenidos en experimentos del estado del arte, es la distancia euclídea.

Estos son los resultados obtenidos para un valor de $k = 7$ (para los valores $k = 1, 3$ y 5 nos quedaríamos con el primero, los tres primeros y los cinco primeros, respectivamente):

Distancia	Imagen	Código IRMA asociado
$1,2469 * 10^3$		1121-220-230-700
$1,3224 * 10^3$		1121-220-230-700
$1,3756 * 10^3$		1121-120-200-700
$1,3893 * 10^3$		1121-220-230-700

3.- Análisis del sistema de clasificación de imágenes

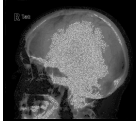

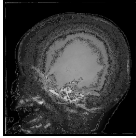
$1,3936 * 10^3$		1121-220-230-700
$1,4222 * 10^3$		1121-120-200-700
$1,4365 * 10^3$		1121-220-230-700

Tabla 20. Resultados algoritmo k-NN de imagen de ejemplo

Como puede apreciarse en vista de los resultados expuestos en la tabla anterior, para un conjunto en el que se incluyen todas las características obtenidas anteriormente aplicándoles un peso de valor 1 a cada una de ellas:

- Para $k = 1$ sólo habría un posible resultado con probabilidad de ocurrencia 1 (ya que se ha obtenido un único vecino), que es 1121-220-230-700. En este caso se obtendría un resultado correcto, ya que la imagen para la que la distancia es menor pertenece a la misma clase que la imagen que se ha elegido como ejemplo del conjunto de test.
- Para $k = 3$ habría dos posibles resultados, con probabilidad de ocurrencia $2/3$ para 1121-220-230-700 y probabilidad de ocurrencia $1/3$ para 1121-120-200-700, de manera que como se elige el resultado con mayor probabilidad de ocurrencia, en este caso el resultado final sería 1121-220-230-700, que al igual que en el caso $k = 1$, también es un resultado correcto.
- Para $k = 5$ habría también dos posibles resultados, con probabilidad de ocurrencia $4/5$ para 1121-220-230-700 y probabilidad de ocurrencia $1/5$ para 1121-120-200-700, de manera que como se elige el resultado con mayor probabilidad de ocurrencia, en este caso elegiríamos 1121-220-230-700, que también sería un resultado correcto.
- Para $k = 7$ se tendrían de nuevo los dos mismos posibles resultados, con probabilidad de ocurrencia $5/7$ para 1121-220-230-700 y probabilidad de ocurrencia $2/7$ para 1121-120-200-700, de manera que como se elige el resultado con mayor probabilidad de ocurrencia, en este caso elegiríamos 1121-220-230-700, que también sería un resultado correcto.

4.- IMPLEMENTACIÓN DEL SISTEMA DE CLASIFICACIÓN DE IMÁGENES

4.1.- Introducción

El diseño del sistema de clasificación de imágenes basado en contenido ha sido implementado en MATLAB, un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio [66].

4.2.- Funciones implementadas

Las funciones que forman el sistema de clasificación de imágenes basado en contenido se pueden dividir en tres bloques fundamentales que se explican a continuación.

4.2.1.- Funciones de procesamiento de imágenes

En primer lugar recordar que tenemos dos conjuntos de imágenes, uno de entrenamiento con imágenes previamente clasificadas, y otro de test con imágenes sin clasificar. Ambos conjuntos de imágenes estarán almacenados en dos directorios distintos.

→ procesaDirectorio

Esta función permite, a partir de la ruta del directorio donde se encuentra el conjunto de imágenes que se quiere procesar, obtener la matriz de características del mismo. En el caso del conjunto de imágenes de entrenamiento, la matriz obtenida constituirá la propia base de datos del sistema de clasificación de imágenes basado en contenido. En el caso del conjunto de imágenes de test, la matriz obtenida será la que contenga las características de este conjunto de imágenes y se usará para la posterior clasificación de las mismas y su inclusión en la base de datos del sistema.

Parámetros de entrada:

- Ruta del directorio: ruta del directorio donde se encuentran las imágenes a procesar.
- Resize X, Resize Y: tamaño al que se redimensionan las imágenes tanto en el eje X como en el eje Y (en píxeles).
- Tamaño X, Tamaño Y: tamaño de las diferentes secciones en que se divide una imagen para obtener sus características locales (en píxeles).

4.- Implementación del sistema de clasificación de imágenes

Parámetros de salida:

- Matriz de características: matriz que contiene todas las características globales y locales de las imágenes de un directorio.
- Vector de posiciones: vector con las posiciones que ocupa cada característica incluida en la matriz teniendo en cuenta sus valores globales y locales.

Esta función recorre las imágenes incluidas en el directorio una a una (en un bucle) y ejecuta la función “*procesalmagen*” explicada a continuación, que nos devuelve los vectores de características de cada una de las imágenes, con los que se forma la matriz.

→ **procesalmagen**

Esta función nos permite, a partir del nombre de la imagen que se está procesando, la obtención del vector de características.

Parámetros de entrada:

- Nombre de la imagen: nombre de la imagen que vamos a procesar.
- Resize X, Resize Y: tamaño al que se redimensionan las imágenes tanto en el eje X como en el eje Y (en píxeles).
- Tamaño X, Tamaño Y: tamaño de las diferentes secciones en que se divide una imagen para obtener sus características locales (en píxeles).

Parámetros de salida:

- Vector de características: vector que contiene todas las características globales y locales de una imagen.
- Vector de posiciones: vector con las posiciones que ocupa cada característica incluida en la matriz teniendo en cuenta sus valores globales y locales.

Además de las características, a este vector de características se le añade una posición al final en la que se almacenará el nombre de la imagen.

Para la obtención de las características se llama a la función “*ExtraeCaracterísticas*” explicada a continuación. Para la obtención de características hay que hacer un tratamiento previo de la imagen, de manera que se obtiene la información de la misma que es la que se pasa a la función propia de extracción de características. El tratamiento que se hace de la imagen es el mismo para todas las características, excepto para los estadísticos y la matriz de co-ocurrencia, que no se hace ningún tratamiento, simplemente se lee la imagen y se redimensiona:

- Lectura de la imagen.
- Redimensión de la imagen.
- Conversión a double de los datos obtenidos de la redimensión de la imagen.
- Ecuilización de los datos obtenidos de la conversión a double.

El formato de la matriz de características obtenida es el siguiente:

$$\begin{bmatrix} c11 & c12 & c13 & \dots & c1M & \text{nombre1} \\ c21 & c22 & c23 & \dots & c2M & \text{nombre2} \\ \dots & \dots & \dots & & \dots & \dots \\ cN1 & cN2 & cN3 & \dots & cNM & \text{nombreN} \end{bmatrix}$$

→ baseDatos

Esta función permite almacenar la clase a la que pertenece cada una de las imágenes del conjunto de entrenamiento, y que como se ha explicado anteriormente, están previamente clasificadas.

Parámetros de entrada:

- Matriz de características de entrenamiento: matriz de características del conjunto de entrenamiento.

Parámetros de salida:

- Nueva matriz de entrenamiento: matriz de características del conjunto de entrenamiento con la clase añadida.
- Nombres de imágenes: vector que contiene todos los nombres de las imágenes.
- Clases de imágenes: vector que contiene todos los códigos (strings de 13 caracteres) de las imágenes. El índice de este vector se corresponde con el índice del vector de nombres.
- Clases: vector en el que se almacenan las diferentes clases con las que se va a trabajar. El índice dentro de este vector es el que se utilizará como identificador de la clase y el dato que se almacenará en la matriz de características.

4.2.2.- Funciones de extracción de características

Como se ha explicado anteriormente, para la creación de la base de datos y la posterior clasificación de las imágenes se hace uso de características obtenidas de las imágenes tanto de forma global como de forma local. Se han implementado dos funciones para extraer ambos tipos de características.

→ **ExtraeCaracterísticas**

Esta función obtiene las características de una imagen tanto de forma global como de forma local y las almacena en un vector, en el que las va ordenando por características, de manera que la información global y local de cada una va de forma consecutiva.

Parámetros de entrada:

- Información imagen: información de la imagen que se le pasa a las funciones que extraen cada característica.
- Información imagen auxiliar: información de la imagen que se le pasa a las funciones que extraen los estadísticos y la matriz de co-ocurrencia.
- Tamaño X, Tamaño Y: tamaño de las diferentes secciones en que se divide una imagen para obtener sus características locales (en píxeles).

Parámetros de salida:

- Vector de características: vector que contiene todas las características globales y locales de la imagen.
- Vector de posiciones: vector con las posiciones que ocupa cada característica incluida en la matriz teniendo en cuenta sus valores globales y locales.

El formato del vector de características que se obtiene para cada imagen en el siguiente:

$$(aa\ a\dots abb\dots b\ \dots\ zz)$$

Donde cada letra representa la información de una característica, de manera que en las primeras posiciones de esa letra se almacena la característica de forma global, y en las últimas posiciones de esa misma letra, la característica de forma local.

Dentro de esta función se usan las funciones implementadas para la extracción de las características descritas en apartados anteriores. Estas funciones son las que obtienen las características de las imágenes, tanto de manera global como de manera local, y las almacenan en un vector.

→ Histograma

En esta función se obtiene el histograma de la imagen a analizar. Para ello se utiliza una función ya implementada en MATLAB, "*imhist*", a la que se le pasa como parámetro de entrada la información que se obtiene procesar la imagen con los pasos explicados anteriormente, y el número de niveles de gris que se quieran obtener.

Parámetros de entrada:

- Información imagen: información de la imagen después de su procesado.
- Niveles: número de niveles del histograma.

Parámetros de salida:

- Valor histograma: vector con el valor de la frecuencia de cada uno de los niveles de gris (tantos como se han especificado al llamar a la función).

→ Estadísticos

Esta función devuelve el valor de los cinco estadísticos basados en el histograma de primer orden de la imagen a analizar. Para ello se utilizan diferentes funciones implementadas en MATLAB: "mean" para obtener la media, "median" para obtener la mediana, "std" para obtener la desviación típica, a partir de la cual obtendremos la varianza, "skewness" para obtener el coeficiente de asimetría, y por último "kurtosis" para obtener el coeficiente de Curtosis.

Parámetros de entrada:

- Información de la imagen.

Parámetros de salida:

- Estadísticos: vector con el valor de los estadísticos de la imagen.

→ DimensionFractal

Esta función devuelve el valor de la dimensión fractal de la imagen a analizar.

Parámetros de entrada:

- Información imagen: información de la imagen después de su procesado.

4.- Implementación del sistema de clasificación de imágenes

Parámetros de salida:

- Dimensión fractal: valor de la dimensión fractal de la imagen.

→ Gabor

Esta función devuelve el valor de la media y la varianza de las características de Gabor obtenidas de la imagen a analizar.

Parámetros de entrada:

- Información imagen: información de la imagen después de su procesado.
- Escalas: número de escalas.
- Orientaciones: número de orientaciones.

Parámetros de salida:

- Estadísticos Gabor: vector con el valor de los estadísticos de la imagen después de pasar por los filtros de Gabor.

Esta función hace uso de la función “*Gaborconvolve*”, que se explica a continuación:

→ Gaborconvolve

Esta función realiza la convolución de la imagen con cada uno de los filtros que se obtienen de las posibles combinaciones de escalas y orientaciones.

Parámetros de entrada:

- Información imagen: información de la imagen después de su procesado.
- Número de escalas que se van a estudiar.
- Número de orientaciones que se van a estudiar.
- Valor de la longitud de onda el filtro de mínima escala. En este caso se usará el valor 3.
- Factor multiplicador entre filtros sucesivos. En este caso se usará el valor 2.
- Ratio de la desviación típica de la Gaussiana que describe la función de transferencia del filtro de Gabor. En este caso se usará el valor 0.65.

4.- Implementación del sistema de clasificación de imágenes

- Ratio del intervalo angular entre las orientaciones de los filtros y la desviación típica de la función angular Gaussiana. En este caso se usará el valor 1.5.

Parámetros de salida:

- Imagen filtrada: información de la imagen filtrada por el correspondiente filtro de Gabor definido por una escala y una orientación.

→ Laws

Esta función devuelve el valor de la energía de la imagen a analizar una vez que se le aplican los filtros de Laws. En este caso se han combinado los filtros de Laws que obtienen características similares.

Parámetros de entrada:

- Información imagen: información de la imagen después de su procesado.

Parámetros de salida:

- Energías Laws: vector con el valor de las energías de la imagen después de pasar por los filtros de Laws.

Hace uso de las funciones “*FiltrosLaws*” y “*Energy*” que se explican a continuación.

→ FiltrosLaws

Esta función filtra la imagen original con la máscara de filtro de Laws correspondiente.

Parámetros de entrada:

- Información imagen: información de la imagen después de su procesado.
- ID filtro: identificador del filtro a aplicar.

Parámetros de salida:

- Imagen filtrada: información de la imagen después del filtro.

→ Energy

Esta función devuelve, a partir de la matriz de información de una imagen (en este caso una imagen filtrada con la máscara del filtro de Laws correspondiente), la energía de la misma.

Parámetros de entrada:

- Información imagen: matriz con la información de la imagen después del filtro.

Parámetros de salida:

- Energía: valor de la energía de la imagen después del filtro.

→ Tamura

Esta función obtiene las características de Tamura de la imagen a analizar.

Parámetros de entrada:

- Información imagen: matriz con la información de la imagen después de su procesado.

Parámetros de salida:

- Estadísticos: valor de los estadísticos de Tamura.

Hace uso de las funciones “*coarseness*”, “*contraste*” y “*directionality*” explicadas a continuación.

→ Coarseness

Esta función devuelve la “tosquedad” de la imagen.

Parámetros de entrada:

- Información imagen: matriz con la información de la imagen después de su procesado.

4.- Implementación del sistema de clasificación de imágenes

Parámetros de salida:

- Tosquedad: valor de la tosquedad de la imagen.

→ Contraste

Esta función devuelve el contraste de la imagen.

Parámetros de entrada:

- Información imagen: matriz con la información de la imagen después de su procesado.

Parámetros de salida:

- Contraste: valor del contraste de la imagen.

→ Directionality

Esta función devuelve la direccionalidad de la imagen.

Parámetros de entrada:

- Información imagen: matriz con la información de la imagen después de su procesado.

Parámetros de salida:

- Direccionalidad: valor de la direccionalidad de la imagen.

→ MatrizCoocurrencia

Esta función obtiene los estadísticos de la matriz de co-ocurrencia (o *Gray Level Co-occurrence Matrix*) obtenida a partir de la imagen a analizar. Hace uso de las funciones “*GLCM*”, “*homogeneity*”, “*entropy*”, “*energy*”, “*contrast*” y “*correlación*” explicadas a continuación:

Parámetros de entrada:

- Información imagen: matriz con la información de la imagen después de su procesado.

4.- Implementación del sistema de clasificación de imágenes

Parámetros de salida:

- Estadísticos: estadísticos de la matriz de co-ocurrencia de la imagen a analizar.

→ GLCM

Esta función devuelve la matriz de co-ocurrencia de la imagen a analizar. A parte de la información de la imagen de la que se quiere obtener la matriz, hay que pasarle como parámetros de entrada la distancia entre píxeles que se va a tomar, y un vector de orientaciones para las que se calculará la matriz.

Parámetros de entrada:

- Información imagen: matriz con la información de la imagen después de su procesado.
- Distancia: distancia que se tendrá en cuenta entre píxeles.
- Orientaciones: vector con las orientaciones para las que se calculará la matriz.

Parámetros de salida:

- Matriz de co-ocurrencia: matriz de co-ocurrencia de la imagen.

→ Homogeneity

Esta función devuelve la homogeneidad de la matriz de co-ocurrencia.

Parámetros de entrada:

- Matriz de co-ocurrencia: matriz de co-ocurrencia de la imagen.

Parámetros de salida:

- Homogeneidad: homogeneidad de la matriz de co-ocurrencia.

→ Entropy

Esta función devuelve la entropía de la matriz de co-ocurrencia.

4.- Implementación del sistema de clasificación de imágenes

Parámetros de entrada:

- Matriz de co-ocurrencia: matriz de co-ocurrencia de la imagen.

Parámetros de salida:

- Entropía: entropía de la matriz de co-ocurrencia.

→ Energy

Esta función devuelve la energía de la matriz de co-ocurrencia.

Parámetros de entrada:

- Matriz de co-ocurrencia: matriz de co-ocurrencia de la imagen.

Parámetros de salida:

- Energía: energía de la matriz de co-ocurrencia.

→ Contrast

Esta función devuelve el contraste de la matriz de co-ocurrencia.

Parámetros de entrada:

- Matriz de co-ocurrencia: matriz de co-ocurrencia de la imagen.

Parámetros de salida:

- Contraste: contraste de la matriz de co-ocurrencia.

→ Correlacion

Esta función devuelve la correlación de la matriz de co-ocurrencia.

Parámetros de entrada:

- Matriz de co-ocurrencia: matriz de co-ocurrencia de la imagen.

4.- Implementación del sistema de clasificación de imágenes

Parámetros de salida:

- Correlación: correlación de la matriz de co-ocurrencia.

→ CoeficientesDCT

Esta función devuelve los coeficientes de la transformada discreta del coseno de la imagen a analizar. Para ello hace uso de la función ya implementada en MATLAB “*dct*”, que devuelve los coeficientes, que posteriormente serán normalizados.

Parámetros de entrada:

- Información imagen: matriz con la información de la imagen después de su procesado.

Parámetros de salida:

- Coeficientes DCT: coeficientes de la transformada discreta del coseno de la imagen a analizar. Sólo se cogen los coeficientes que interesan, que son los que están asociados con las frecuencias altas.

→ CoeficientesDWT

Esta función devuelve la media y la desviación típica de los coeficientes de la transformada de Wavelet discreta, en cada una de las tres bandas, de la imagen a analizar. Para ello hace uso de la función ya implementada en MATLAB “*dwt*”, que devuelve los coeficientes.

Parámetros de entrada:

- Información imagen: matriz con la información de la imagen después de su procesado.

Parámetros de salida:

- Coeficientes DWT: valores de la media y la varianza de los coeficientes de la transformada discreta de Wavelet de la imagen a analizar.

→ CoeficientesDFT

Esta función devuelve los coeficientes de la transformada discreta de Fourier de la imagen a analizar. Para ello hace uso de la función ya implementada en MATLAB “*fft*”, que devuelve los coeficientes, que posteriormente serán normalizados y obtenido el módulo.

Parámetros de entrada:

- Información imagen: matriz con la información de la imagen después de su procesado.

Parámetros de salida:

- Coeficientes DFT: coeficientes de la transformada discreta de Fourier de la imagen a analizar. Se calcula el módulo de los mismos.

4.2.3.- Funciones de clasificación

Una vez que ya se tienen las matrices de características tanto del conjunto de imágenes de entrenamiento como de test se procede, para cada una de las imágenes de test, a buscar las k distancias menores a los vectores de características de las imágenes de entrenamiento. Lo primero que hay que hacer es calcular las distancias de cada una de las imágenes de test con cada una de las imágenes de entrenamiento. Esto se hace mediante las funciones explicadas a continuación.

→ distancias

Esta función obtiene una matriz en la que se almacena, para cada imagen de test, la distancia a cada una de las imágenes de entrenamiento.

Parámetros de entrada:

- Matriz de características de entrenamiento: sin el nombre de la imagen ni su clase asociada.
- Matriz de características de test: sin el nombre de la imagen.

Parámetros de salida:

- Matriz de distancias: matriz que almacena la distancia de cada una de las imágenes de test a cada una de las imágenes de entrenamiento. Tiene por

4.- Implementación del sistema de clasificación de imágenes

tanto, tantas filas como imágenes haya en el conjunto de test, y tantas columnas como imágenes haya en el conjunto de entrenamiento.

El formato de la matriz de distancias es el siguiente:

$$\begin{pmatrix} \text{dis tan cia1_1} & \text{dis tan cia1_2} & \dots & \text{dis tan cia1_K} \\ \text{dis tan cia2_1} & \text{dis tan cia2_2} & \dots & \text{dis tan cia2_K} \\ \dots & \dots & \dots & \dots \\ \text{dis tan ciaN_1} & \text{dis tan ciaN_2} & \dots & \text{dis tan ciaN_K} \end{pmatrix}$$

→ distanciasMin

Esta función obtiene una matriz en la que se almacena, para cada imagen de test, las k distancias mínimas, junto con el nombre y la clase a la que pertenecen las imágenes de entrenamiento a las que corresponden estas distancias mínimas.

Parámetros de entrada:

- Matriz de distancias: matriz que almacena la distancia de cada una de las imágenes de test a cada una de las imágenes de entrenamiento.
- Matriz de características de entrenamiento: matriz de características de entrenamiento, de donde obtenemos su nombre y su clase asociada.
- k: número de vecinos más cercanos obtenidos.

Parámetros de salida:

- Matriz de distancias mínimas: matriz que almacena las k distancias mínimas a k imágenes de entrenamiento, junto con su nombre y su clase asociada. Tiene por tanto, tantas filas como imágenes haya en el conjunto de test, y tantas 3 *k columnas.

El formato de la matriz de distancias es el siguiente:

$$\begin{pmatrix} \text{dis tan cia1_1} & \text{clase1_1} & \text{nombre1_1} & \text{dis tan cia1_2} & \text{clase1_2} & \text{nombre1_2} & \dots & \text{dis tan cia1_K} & \text{clase1_K} & \text{nombre1_k} \\ \text{dis tan cia2_1} & \text{clase2_1} & \text{nombre2_1} & \text{dis tan cia2_2} & \text{clase2_2} & \text{nombre2_2} & \dots & \text{dis tan cia2_K} & \text{clase1_K} & \text{nombre2_k} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \text{dis tan ciaN_1} & \text{claseN_1} & \text{nombreN_1} & \text{dis tan ciaN_2} & \text{claseN_2} & \text{nombreN_2} & \dots & \text{dis tan ciaN_K} & \text{clase1_K} & \text{nombreN_k} \end{pmatrix}$$

Una vez que ya se tiene la matriz con los k vecinos más cercanos y sus distancias se busca el código de la imagen. Para ellos se implementa la función que se explica a continuación.

→ clasificación

A partir de la matriz de distancias mínimas se obtiene la probabilidad de aparición de cada clase y se asigna a cada imagen de test la clase más probable. En caso de empate se elige la clase de distancia menor.

Parámetros de entrada:

- Matriz de distancias mínimas: matriz que almacena las k distancias mínimas a k imágenes de entrenamiento, junto con su nombre y su clase asociada. Tiene por tanto, tantas filas como imágenes haya en el conjunto de test, y tantas $3 * k$ columnas.
- Matriz de características de test: matriz de características de test, donde se guarda la clase asignada a cada imagen en la última posición.

Parámetros de salida:

- Matriz de características de test nueva: matriz de características de test con la clase ya asignada.

Esta nueva matriz de características con el código asociado a cada imagen se puede anexar a la matriz de características y códigos del conjunto de imágenes de entrenamiento para ampliar la base de datos y tener más información y próximas búsquedas y clasificaciones.

4.2.4.- Funciones de evaluación de los resultados

Una vez que ya se tienen todas las imágenes del conjunto de test clasificadas hay que evaluar los resultados obtenidos, comparando la clase asignada a cada imagen con la verdadera clase a la que pertenecen. Para ellos se ha implementado la siguiente función.

→errorRate

Calcula la tasa de error, que es la probabilidad de error total, es decir, el número de errores dividido por el número total de imágenes de test, y la puntuación o *score*, que es el número total de aciertos.

Parámetros de entrada:

- Matriz de características de test: matriz de características de test completa, con su clase ya asignada.

4.- Implementación del sistema de clasificación de imágenes

- Nombres de imágenes: vector que contiene todos los nombres de las imágenes.
- Clases de imágenes: vector que contiene todos los códigos (strings de 13 caracteres) de las imágenes. El índice de este vector se corresponde con el índice del vector de nombres.
- Clases: vector en el que se almacenan las diferentes clases con las que vamos a trabajar. El índice dentro de este vector es el que se utilizará como identificador de la clase y el dato que se almacenará en la matriz de características.

Parámetros de salida:

- Tasa de error: tasa de error del sistema para las matrices de características especificadas.

5.- DESARROLLO Y PRUEBAS

5.1.- Introducción

En este apartado se van a describir las pruebas realizadas con el sistema de clasificación automática de imágenes basado en contenido que se ha implementado. Para ello, en primer lugar se ha obtenido la base de datos de entrenamiento, con todas las características explicadas en los capítulos de estado del arte y de diseño del sistema, el nombre de la imagen y su clase asociada (correspondiente con el código IRMA de la imagen). Una vez que se tiene esta matriz se elige una imagen del conjunto de test y se estudia el funcionamiento del sistema para distintas combinaciones de características y distinto número de vecinos en el algoritmo k-NN.

Como se ha explicado en apartados en los capítulos de estado del arte y de diseño del sistema, se va a trabajar con un conjunto de nueve características distintas:

- Histograma
- Estadísticos
- Dimensión fractal
- Características de Gabor
- Características de Tamura
- Matriz de co-ocurrencia
- Transformada discreta del coseno
- Transformada discreta de Wavelet
- Transformada discreta de Fourier

Asignando un peso 1 o 0 a cada una de las características se estudia la importancia o no importancia de cada una de ellas a la hora de clasificar una imagen. Si tenemos en cuenta sólo estos dos valores de pesos se tienen 2^9 (512) posibles combinaciones de características, que es un número de casos demasiado grande como para abarcarlo entero. Se van a estudiar en primer lugar los diez casos básicos, que son, en primer lugar el caso de incluir todas las características, y después nueve casos en los que se elimina una característica en cada uno de ellos. De esta forma se puede ver qué características aportan más información que otras a la hora de decidir, y cuáles de ellas no son imprescindibles debido a que los resultados no empeoran si son eliminadas. Además se incluye el estudio de las características una a una por separado, para poder ver la clase de imágenes que se obtienen con la comparación de cada una de ellas.

Una vez hecho este primer estudio en el que se han encontrado las características más influyentes, se van a incluir otros en los que se asignarán pesos mayores que 1 a las

características que más importancia tienen a la hora de decidir. Sin embargo, estos estudios se incluirán sólo en la evaluación de los resultados del sistema completo, con las 1.000 imágenes del conjunto de test, en el siguiente apartado.

Para cada uno de los análisis descritos en los dos párrafos anteriores se va a incluir un parámetro variable, que será el número de vecinos más cercanos elegido para el algoritmo k-NN, pudiendo observar así cómo, según se aumenta el valor de este parámetro los resultados mejoran, hasta un punto en el que comienzan a empeorar o dejan de mejorar en relación al coste computacional que supone.

En el Apéndice A está incluida la tabla de las diferentes clases a las que pertenecen las imágenes de los conjuntos de test y entrenamiento con los que se ha trabajado, y el número de elementos que contiene cada una. Tal y como se ha explicado en apartados anteriores, se observa cómo para algunas de las clases el número de individuos pertenecientes a las mismas es muy bajo, de manera que elegir un valor de k muy grande no mejora el funcionamiento del algoritmo de clasificación, ya que si este número es mayor que el número de individuos pertenecientes a la clase a la cual debería pertenecer la imagen que se está estudiando, obligatoriamente el algoritmo comenzará a seleccionar como vecinos más cercanos imágenes de clases incorrectas, y según se aumente el valor de este número de vecinos, la probabilidad de aparición de la clase correcta disminuirá hasta el punto de que sea menor que la probabilidad del resto de clases y, por tanto, se tomará una decisión incorrecta.

5.2.- Pruebas realizadas

La imagen seleccionada para la realización de los diferentes análisis, elegida de manera aleatoria dentro de una clase en la que tanto en el conjunto de entrenamiento como en el de test existe un número grande de individuos, que pertenece al conjunto de imágenes de test, es la siguiente.



Código IRMA asociado: 1123-110-500-000

Figura 15. Imagen ejemplo para pruebas

5.2.1.- Características por separado

En este apartado se estudia el resultado de la clasificación de la imagen de ejemplo usando las características una por una de manera individual. Los siguientes resultados mostrados son los resultados obtenidos para un valor de $k = 7$ (para los valores $k = 1, 3$ y 5 se quedarían el primero, los tres primeros y los cinco primeros, respectivamente).

- **Histograma**

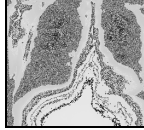
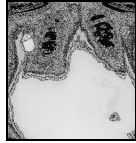
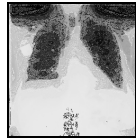
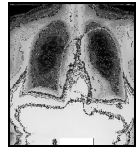

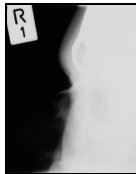

Distancia	Imagen	Código IRMA asociado
5,1112		1123-110-500-000
5,1116		1123-110-500-000
5,1187		1123-110-500-000
5,1196		1123-110-500-000
5,12		1121-120-514-700
5,1205		1121-220-213-700
5,1211		1121-210-310-700

Tabla 21. Resultados de clasificación para el histograma

- Resultado para k = 1: 1123-110-500-000 ✓
- Resultado para k = 3: 1123-110-500-000 ✓
- Resultado para k = 5: 1123-110-500-000 ✓
- Resultado para k = 7: 1123-110-500-000 ✓

- Estadísticos

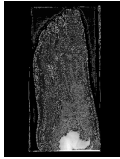



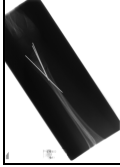

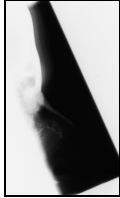
Distancia	Imagen	Código IRMA asociado
0,8924		1121-490-915-700
1,9838		1121-220-213-700
1,9838		1121-200-411-700
1,9838		1123-211-500-000
1,9838		1121-240-433-700
1,9838		1121-240-433-700
1,9838		1121-210-213-700

Tabla 22. Resultados de clasificación para los estadísticos

- Resultado para k = 1: 1121-490-915-700 ✖
- Resultado para k = 3: 1121-490-915-700 ✖
- Resultado para k = 5: 1121-490-915-700 ✖
- Resultado para k = 7: 1121-240-433-700 ✖

- Dimensión fractal


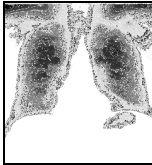
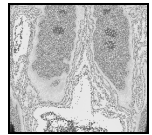
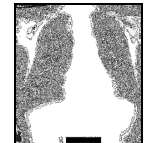

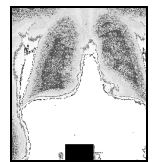
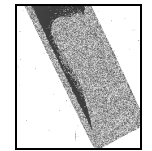
Distancia	Imagen	Código IRMA asociado
0,0255		1123-110-500-000
0,0304		1123-110-500-000
0,0321		1123-110-500-000
0,0345		1123-110-500-000
0,0354		1123-110-500-000
0,0356		1123-110-500-000
0,0375		1121-120-911-700

Tabla 23. Resultados de clasificación para la dimensión fractal

- Resultado para $k = 1$: 1123-110-500-000 ✓
- Resultado para $k = 3$: 1123-110-500-000 ✓
- Resultado para $k = 5$: 1123-110-500-000 ✓
- Resultado para $k = 7$: 1123-110-500-000 ✓

- Características de Gabor




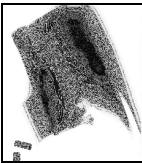
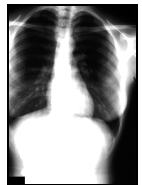


Distancia	Imagen	Código IRMA asociado
2,6782		1123-211-500-000
2,7367		1123-211-500-000
2,7630		1123-211-500-000
2,7936		1121-228-310-700
2,8071		1123-110-500-000
2,8115		1121-200-411-700
2,8203		1123-211-500-000

Tabla 24. Resultados de clasificación para las características de Gabor

- Resultado para k = 1: 1123-211-500-000 ❌
- Resultado para k = 3: 1123-211-500-000 ❌
- Resultado para k = 5: 1123-211-500-000 ❌
- Resultado para k = 7: 1123-211-500-000 ❌

- Características de Tamura

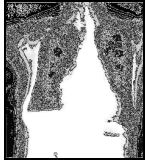


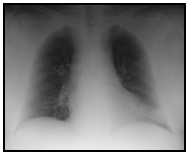
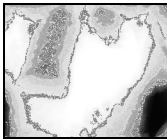
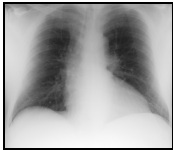

Distancia	Imagen	Código IRMA asociado
0,2188		1123-110-500-000
0,2234		1121-120-961-700
0,2248		1123-110-500-000
0,2365		1123-127-500-000
0,2404		1123-127-500-000
0,2442		1123-127-500-000
0,2484		1123-110-500-000

Tabla 25. Resultados de clasificación para las características de Tamura

- Resultado para k = 1: 1123-110-500-000 ✓
- Resultado para k = 3: 1123-110-500-000 ✓
- Resultado para k = 5: 1123-110-500-000 ✓
- Resultado para k = 7: 1123-110-500-000 ✓

- Matriz de co-ocurrencia

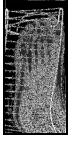


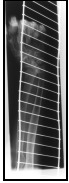

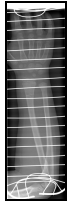
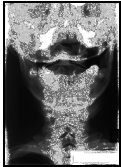
Distancia	Imagen	Código IRMA asociado
4,9517170		1121-490-915-700
4,9517231		1121-120-310-700
4,9517234		1121-120-930-700
4,9517235		1121-120-434-700
4,9517344		1121-110-414-700
4,9517373		1121-120-438-700
4,9517375		1121-120-310-700

Tabla 26. Resultados de clasificación para la matriz de co-ocurrencia

- Resultado para k = 1: 1121-490-915-700 ✖
- Resultado para k = 3: 1121-490-915-700 ✖
- Resultado para k = 5: 1121-490-915-700 ✖
- Resultado para k = 7: 1121-120-310-700 ✖

- Transformada discreta del coseno, DCT

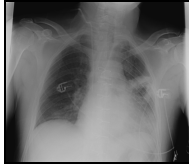
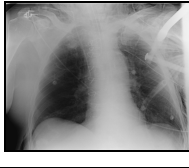

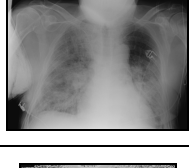
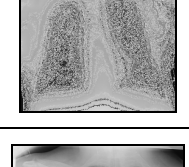
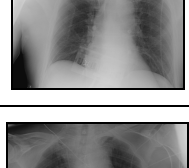
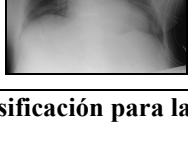
Distancia	Imagen	Código IRMA asociado
0,1698		1123-127-500-000
0,1752		1123-127-500-000
0,1960		1123-127-500-000
0,2107		1123-127-500-000
0,2442		1123-127-500-000
0,2446		1123-127-500-000
0,2488		1123-127-500-000

Tabla 27. Resultados de clasificación para la transformada discreta del coseno, DCT

- Resultado para k = 1: 1123-110-500-000 ✓
- Resultado para k = 3: 1123-110-500-000 ✓
- Resultado para k = 5: 1123-110-500-000 ✓
- Resultado para k = 7: 1123-110-500-000 ✓

- Transformada discreta de Wavelet, DWT



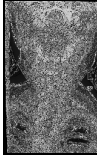

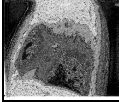
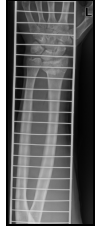
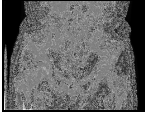
Distancia	Imagen	Código IRMA asociado
0,5161		1121-120-310-700
0,5875		1121-110-414-700
0,5951		1121-120-310-700
0,6024		1121-240-421-700
0,6041		1123-211-500-000
0,6557		1121-120-433-700
0,6734		1121-120-800-700

Tabla 28. Resultados de clasificación para la transformada discreta de Wavelet, DWT

- Resultado para k = 1: 1121-120-310-700 ✖
- Resultado para k = 3: 1121-120-310-700 ✖
- Resultado para k = 5: 1121-120-310-700 ✖
- Resultado para k = 7: 1121-120-310-700 ✖

- Transformada discreta de Fourier, DFT


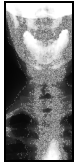

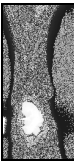
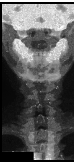


Distancia	Imagen	Código IRMA asociado
2,2740		1121-120-422-700
2,2752		1121-120-310-700
2,2796		1121-110-414-700
2,2834		1121-120-922-700
2,2937		1121-120-310-700
2,2948		1121-120-310-700
2,2969		1121-120-310-700

Tabla 29. Resultados de clasificación para la transformada discreta de Fourier, DFT

- Resultado para k = 1: 1121-120-422-700 ✖
- Resultado para k = 3: 1121-120-422-700 ✖
- Resultado para k = 5: 1121-120-310-700 ✖
- Resultado para k = 7: 1121-120-310-700 ✖

5.2.2.- Inclusión de todas las características

En este apartado se estudia la inclusión de todas las características a la hora de tomar una decisión sobre la clasificación de la imagen.


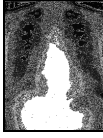
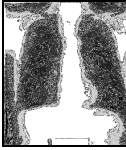

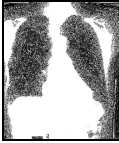


Distancia	Imagen	Código IRMA asociado
8,4459		1123-110-500-000
8,4665		1123-110-500-000
8,5216		1123-110-500-000
8,5322		1121-200-411-700
8,5352		1123-110-500-000
8,54		1123-211-500-000
8,5830		1121-490-415-700

Tabla 30. Resultados de clasificación para todas las características

- Resultado para k = 1: 1123-110-500-000 ✓
- Resultado para k = 3: 1123-110-500-000 ✓
- Resultado para k = 5: 1123-110-500-000 ✓
- Resultado para k = 7: 1123-110-500-000 ✓

5.2.3.- Eliminación de características una por una

En este apartado se estudia la eliminación de cada una de las características, una por una, a la hora de tomar una decisión sobre la clasificación de la imagen, en nueve casos de estudio diferentes (uno por cada característica eliminada). Estos son los resultados obtenidos para un valor de $k = 7$ (para los valores $k = 1, 3$ y 5 se quedarían el primero, los tres primeros y los cinco primeros, respectivamente):

- **Histograma**


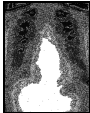





Distancia	Imagen	Código IRMA asociado
6,6711		1123-127-500-000
6,7063		1123-127-500-000
6,7689		1123-127-500-000
6,7798		1121-200-411-700
6,7805		1123-127-500-000
6,7927		1123-211-500-000
6,8376		1121-490-415-700

Tabla 31. Resultados de clasificación eliminando el histograma

- Resultado para $k = 1$: 1123-127-500-000 ✖
- Resultado para $k = 3$: 1123-127-500-000 ✖
- Resultado para $k = 5$: 1123-127-500-000 ✖
- Resultado para $k = 7$: 1123-127-500-000 ✖

- Estadísticos


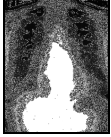





Distancia	Imagen	Código IRMA asociado
8,2096		1123-127-500-000
8,2308		1123-127-500-000
8,2875		1123-127-500-000
8,2984		1121-200-411-700
8,3015		1123-127-500-000
8,3064		1123-211-500-000
8,3506		1121-490-415-700

Tabla 32. Resultados de clasificación eliminando los estadísticos

- Resultado para k = 1: 1123-110-500-000 ✓
- Resultado para k = 3: 1123-110-500-000 ✓
- Resultado para k = 5: 1123-110-500-000 ✓
- Resultado para k = 7: 1123-110-500-000 ✓

- Dimensión fractal


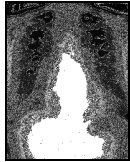
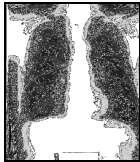




Distancia	Imagen	Código IRMA asociado
8,4453		1123-127-500-000
8,4659		1123-127-500-000
8,5214		1123-127-500-000
8,5313		1121-200-411-700
8,5350		1123-127-500-000
8,5393		1123-211-500-000
8,5809		1121-490-415-700

Tabla 33. Resultados de clasificación eliminando la dimensión fractal

- Resultado para k = 1: 1123-110-500-000 ✓
- Resultado para k = 3: 1123-110-500-000 ✓
- Resultado para k = 5: 1123-110-500-000 ✓
- Resultado para k = 7: 1123-110-500-000 ✓

- Características de Gabor


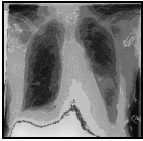

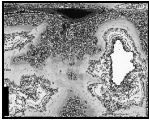
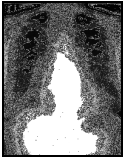
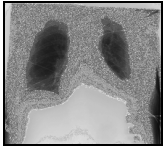
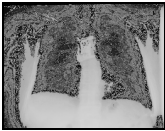
Distancia	Imagen	Código IRMA asociado
7,8902		1121-120-437-700
7,9006		1123-127-500-000
7,9035		1121-490-915-700
7,9070		1121-129-700-400
7,9095		1123-127-500-000
7,9097		1123-127-500-000
7,9105		1123-127-500-000

Tabla 34. Resultados de clasificación eliminando las características de Gabor

- Resultado para k = 1: 1121-120-437-700 ❌
- Resultado para k = 3: 1121-120-437-700 ❌
- Resultado para k = 5: 1123-127-500-000 ❌
- Resultado para k = 7: 1123-127-500-000 ❌

- Características de Tamura


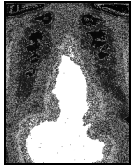




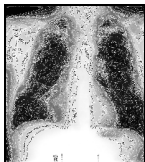
Distancia	Imagen	Código IRMA asociado
8,4408		1123-127-500-000
8,4534		1123-127-500-000
8,5084		1123-127-500-000
8,5175		1123-127-500-000
8,5222		1121-200-411-700
8,5338		1123-211-500-000
8,5536		1123-127-500-000

Tabla 35. Resultados de clasificación eliminando las características de Tamura

- Resultado para k = 1: 1121-120-437-700 ✖
- Resultado para k = 3: 1121-120-437-700 ✖
- Resultado para k = 5: 1123-127-500-000 ✖
- Resultado para k = 7: 1123-127-500-000 ✖

- Matriz de co-ocurrencia


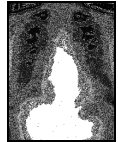
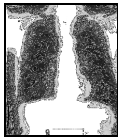

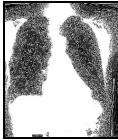


Distancia	Imagen	Código IRMA asociado
6,8419		1123-127-500-000
6,8674		1123-127-500-000
6,9352		1123-127-500-000
6,9482		1121-200-411-700
6,9519		1123-127-500-000
6,9578		1123-211-500-000
7,0105		1121-490-415-700

Tabla 36. Resultados de clasificación eliminando la matriz de co-ocurrencia

- Resultado para k = 1: 1121-120-437-700 ✖
- Resultado para k = 3: 1121-120-437-700 ✖
- Resultado para k = 5: 1123-127-500-000 ✖
- Resultado para k = 7: 1123-127-500-000 ✖

- Transformada discreta del coseno, DCT



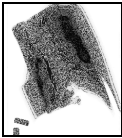

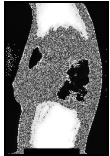
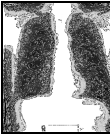

Distancia	Imagen	Código IRMA asociado
8,4173		1121-200-411-700
8,4266		1123-127-500-000
8,4287		1121-228-310-700
8,4295		1123-211-500-000
8,4302		1123-211-500-000
8,4415		1123-127-500-000
8,4596		1123-211-500-000

Tabla 37. Resultados de clasificación eliminando la transformada discreta del coseno, DCT

- Resultado para k = 1: 1121-200-411-700 ✖
- Resultado para k = 3: 1121-200-411-700 ✖
- Resultado para k = 5: 1123-211-500-000 ✖
- Resultado para k = 7: 1123-211-500-000 ✖

- Transformada discreta de Wavelet, DWT


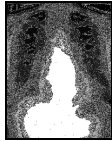

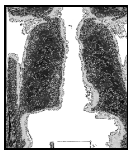
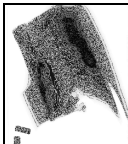
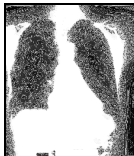

Distancia	Imagen	Código IRMA asociado
8,3408		1123-127-500-000
8,3736		1123-127-500-000
8,4459		1123-211-500-000
8,4496		1123-127-500-000
8,4521		1121-200-411-700
8,4726		1123-127-500-000
8,4910		1123-127-500-000

Tabla 38. Resultados de clasificación eliminando la transformada discreta de Wavelet, DWT

- Resultado para k = 1: 1123-127-500-000 ✖
- Resultado para k = 3: 1123-127-500-000 ✖
- Resultado para k = 5: 1123-127-500-000 ✖
- Resultado para k = 7: 1123-127-500-000 ✖

- Transformada discreta de Fourier, DFT


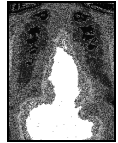
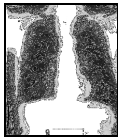

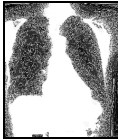


Distancia	Imagen	Código IRMA asociado
8,0839		1123-127-500-000
8,1332		1123-127-500-000
8,1402		1123-127-500-000
8,1434		1121-200-411-700
8,1662		1123-127-500-000
8,1879		1123-211-500-000
8,2247		1121-490-415-700

Tabla 39. Resultados de clasificación eliminando la transformada discreta de Fourier, DFT

- Resultado para k = 1: 1123-127-500-000 ✖
- Resultado para k = 3: 1123-127-500-000 ✖
- Resultado para k = 5: 1123-127-500-000 ✖
- Resultado para k = 7: 1123-127-500-000 ✖

5.3.- Conclusiones ante las pruebas realizadas

Gracias a las pruebas realizadas en este apartado se obtienen algunas conclusiones, aunque se ha de tener en cuenta que estas pruebas han sido realizadas con una sola imagen del conjunto de imágenes de test, y que por tanto no son conclusiones del funcionamiento general del sistema, sino una aproximación a las mismas con el estudio de un individuo.

Además de las pruebas desarrolladas en este apartado de la memoria, se han realizado más con otras imágenes de test, que no se han incluido ya que el tamaño del documento aumentaría considerablemente sin aportar una cantidad de información adicional importante. Sin embargo estas pruebas han llevado a algunas conclusiones que se exponen a continuación con algunos ejemplos que permiten conocer las causas de error más comunes (estas pruebas se han realizado usando todas las características con un peso de valor 1 en todos los casos y para un valor $k = 5$).

→ Imágenes variables dentro de un mismo código

Dentro de una misma clase pueden aparecer imágenes que tengan características que disten muchos en cada caso, y por ese motivo, al compararlas, la distancia va a ser mayor que lo que debería ser para imágenes que tienen el mismo código asociado.

Existen muchos ejemplos que cumplen esta característica, a continuación se muestran algunos casos:

- Código 1121-115-700-400



Figura 16. Variabilidad en imágenes del código 1121-115-700-400

Del conjunto de imágenes de esta clase pertenecientes al conjunto de imágenes de test han sido clasificadas mal 12 de un total de 13 imágenes.

- Código 1121-120-463-700



Figura 17. Variabilidad en imágenes del código 1121-120-463-700

Del conjunto de imágenes de esta clase pertenecientes al conjunto de imágenes de test han sido clasificadas mal 5 de un total de 8 imágenes.

→ Error en el eje de dirección

En muchas ocasiones, el simple hecho de cambiar la dirección de una imagen no implica un gran cambio en el valor de las características, y sin embargo, las imágenes pertenecen a clases diferentes.

Este problema es muy grande si sólo se estudian característica de forma global. En este sistema se ha añadido el estudio de las imágenes de forma local dividiéndolas en 4 partes iguales, de manera que permite comparar las imágenes por partes, pero aún así se sigue dando este problema con imágenes que están referidas a la misma parte del cuerpo pero con distinta orientación. Algunos ejemplos se muestran a continuación.

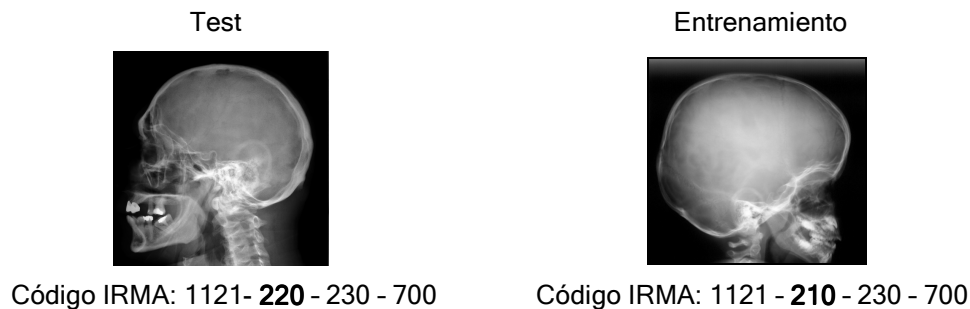


Figura 18. Error en el eje de dirección

→ Error en el eje anatómico

Para las partes pares del cuerpo humano, que tienen un miembro derecho y otro izquierdo, como pueden ser manos, pies, brazos y piernas, existe la posibilidad de equivocación, de la misma manera que sucedía con los cambios de orientación. Aunque como se ha explicado antes, este problema se solventa en gran medida con las características locales, aún se tienen algunos ejemplos en los que se comete este error:

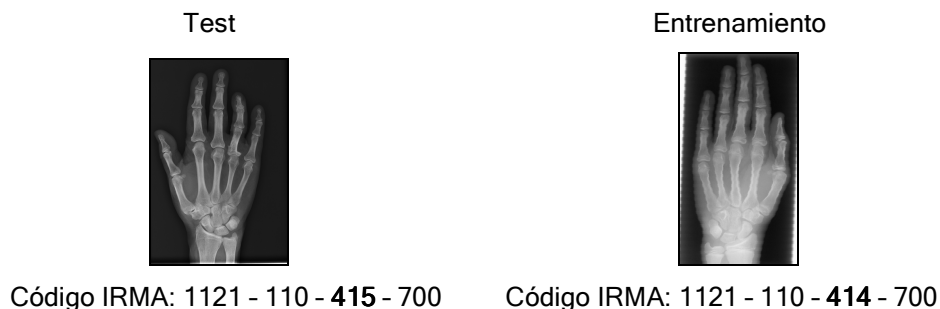


Figura 19. Error en el eje anatómico

→ Error en imágenes giradas

Hay imágenes en la base de datos que no se obtienen de manera correcta, ya que están bien cortadas, bien giradas, y tienen las esquinas blancas por el descuadre de la imagen. Por este motivo, aunque las imágenes pertenezcan a clases totalmente diferentes, las características locales de las mismas coincidirán debido a estas partes blancas que se encuentran en las esquinas, y serán elegidas como vecinos más cercanos de la imagen a analizar provocando así resultados erróneos. Algunos ejemplos se muestran a continuación:

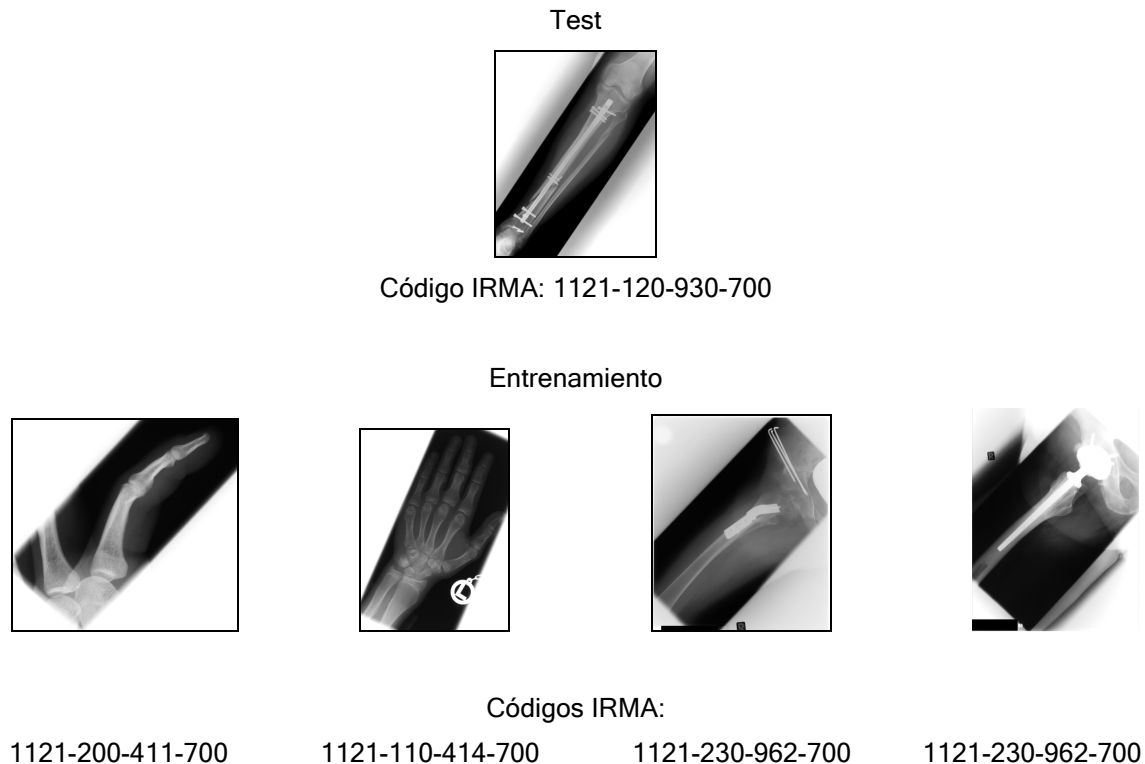


Figura 20. Error en imágenes giradas

6.- EVALUACIÓN DE RESULTADOS

6.1.- Introducción

Una vez realizados los diferentes análisis para una imagen de test en el apartado anterior, en este se van a exponer los resultados obtenidos para cada uno de ellos con la totalidad de imágenes pertenecientes al conjunto de imágenes de test (1.000 imágenes) a través de la obtención de dos indicadores de eficiencia del algoritmo, la puntuación o *score* y la tasa de error, que se explican a continuación.

6.2.- Tasa de error

La tasa de error es el porcentaje de fallos en la totalidad del conjunto de imágenes estudiadas. Para la evaluación del funcionamiento del sistema se ha considerado como correcto la asignación de un código completo de manera satisfactoria, sin dividir el cálculo de este indicador en cuatro sub-tasas de error, una para cada eje calculada independientemente (podría haberse hecho así ya que los ejes son independientes entre sí). Los valores obtenidos están expresados en tanto por uno. Los resultados obtenidos son los mostrados a continuación:

6.2.1.- Características por separado

En la siguiente tabla se muestra la tasa de error obtenida para las 1.000 imágenes de test utilizando sólo una de las características obtenidas, para los cuatro valores de *k* estudiados.

Característica	k = 1	k = 3	k = 5	k = 7
Histograma	0,542	0,542	0,535	0,541
Estadísticos	0,999	0,999	0,998	1
Dimensión fractal	0,696	0,671	0,647	0,633
Características de Gabor	0,408	0,396	0,376	0,382
Características de Tamura	0,657	0,648	0,615	0,612
Matriz de co-ocurrencia	0,994	0,994	0,989	0,975
Transformada discreta del coseno	0,726	0,710	0,705	0,694
Transformada discreta de Wavelet	0,752	0,724	0,706	0,696
Transformada discreta de Fourier	0,853	0,846	0,832	0,829

Tabla 40. Tasa de error de cada característica para distintos valores de *k*

En la siguiente gráfica se muestran los datos de la tabla anterior, de manera que se comparan los resultados de cada uno de los análisis mediante la superposición de las gráficas y su evolución para los distintos valores de *k*.

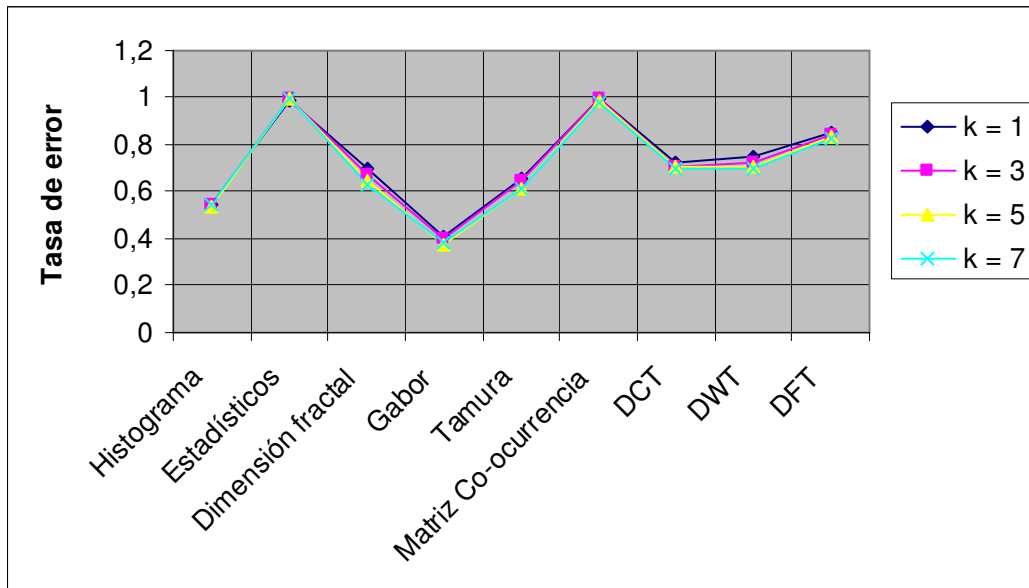


Figura 21. Representación de la tasa de error de cada característica para distintos valores de k

En base a los resultados obtenidos para cada una de las características se puede afirmar, tal y como se había explicado en capítulos anteriores, que hay características que aportan más información que otras para el tipo de imágenes con el que se ha trabajado. En este caso hay dos características que presentan mejor resultado que el resto, que son las características de Gabor y el histograma, mientras que las que peor resultado ofrecen son los estadísticos y la matriz de co-ocurrencia.

Se observa además que los valores obtenidos para los diferentes valores de k son bastante parecidos, aunque se puede apreciar cómo para k = 5 y k = 7 siempre hay una pequeña diferencia y se obtienen resultados ligeramente mejores que para k = 1 y k = 3.

6.2.2.- Inclusión de todas las características

En la siguiente tabla se muestra la tasa de error obtenida para las 1.000 imágenes de test utilizando todas las características, para los cuatro valores de k estudiados.

k = 1	k = 3	k = 5	k = 7
0,384	0,378	0,379	0,386

Tabla 41. Tasa de error obtenida con todas las características y diferentes valores de k

En la siguiente figura se ve una representación gráfica de la tabla anterior:

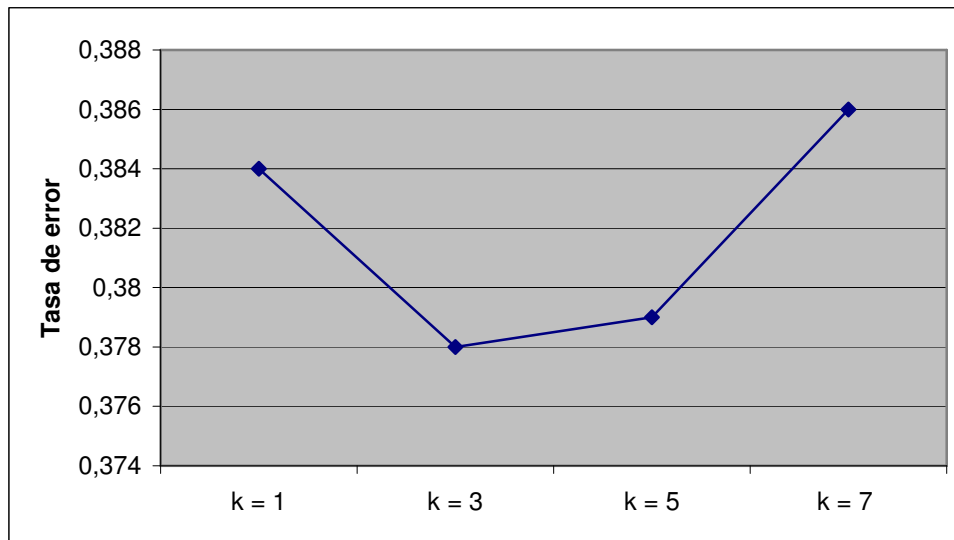


Figura 22. Representación de la tasa de error para todas las características y distintos valores de k

En este caso el mejor resultado se obtiene para $k = 3$, siendo muy similar el resultado obtenido para $k = 5$. Sin embargo se puede observar cómo, tanto para valores bajos ($k = 1$) como alto ($k = 7$) del número de vecinos más cercanos los resultados empeoran, tal y como se ha explicado anteriormente cuando se ha hablado del compromiso a alcanzar para este valor. En la práctica los resultados para los diferentes valores de k son prácticamente iguales, de manera que no tienen relevancia estadística.

6.2.3.- Eliminación de características una por una

En la siguiente tabla se muestra la tasa de error obtenida para las 1.000 imágenes de test eliminando una de las características en cada caso, para los cuatro valores de k estudiados.

Característica eliminada	k = 1	k = 3	k = 5	k = 7
Histograma	0,397	0,375	0,381	0,392
Estadísticos	0,384	0,378	0,379	0,386
Dimensión fractal	0,388	0,380	0,375	0,386
Características de Gabor	0,596	0,591	0,584	0,592
Características de Tamura	0,396	0,383	0,384	0,400
Matriz de co-ocurrencia	0,384	0,376	0,379	0,387
Transformada discreta del coseno	0,389	0,377	0,365	0,370
Transformada discreta de Wavelet	0,414	0,394	0,398	0,401
Transformada discreta de Fourier	0,392	0,365	0,368	0,367

Tabla 42. Tasa de error obtenida para la eliminación de una característica y distintos valores de k

En la siguiente figura se representan los resultados mostrados en la tabla anterior.

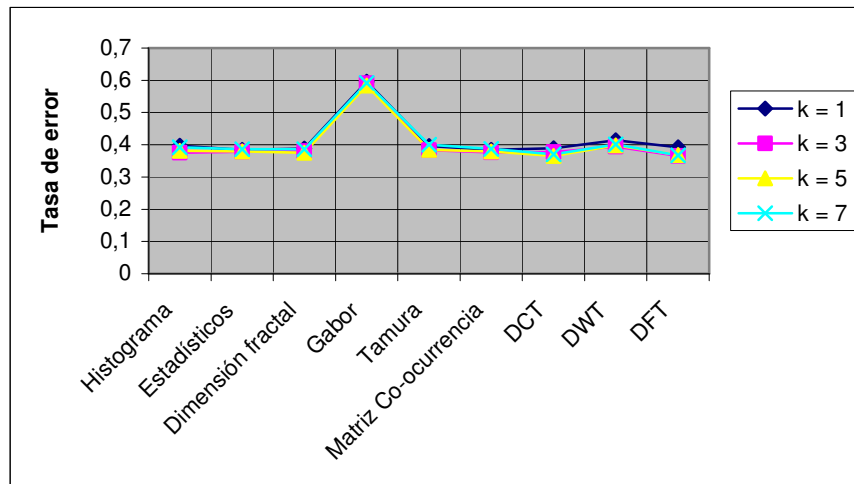


Figura 23. Representación de la tasa de error eliminando una característica y distintos valores de k

Al eliminar una característica se observa el efecto que tiene la inclusión de la misma en el conjunto global de características. En base a los resultados obtenidos se puede afirmar que las características de Gabor son la característica más importante dentro del conjunto, ya que al eliminarla se observa un aumento importante de la tasa de error del sistema.

De la misma manera se observa cómo eliminando la transformada discreta del coseno se obtienen los mejores resultados.

6.2.4.- Asignación de diferentes pesos

Se ha realizado diversos estudios en los que se analiza la eliminación una o varias características y la asignación de pesos con valores distintos a 0 y 1 a las características que, según los estudios realizados en los tres apartados anteriores, son las que más información aportan a la hora de tomar una decisión. No se han incluido en la memoria los resultados de los diferentes estudios realizados ya que la extensión de la memoria sería demasiado grande. De esta manera se han incluido los resultados del estudio que ha ofrecido los mejores resultados. Estos resultados se muestran a continuación.

→ Diseño óptimo:

En este caso se han usado todas las características, pero se ha dado un valor de peso mayor que 1 a las que se ha considerado que más pueden ayudar a la hora de la clasificación, y un peso igual a 1 a las que menos ayudan a esta tarea (en base a los resultados obtenidos en los apartados anteriores).

En la siguiente tabla se muestran los pesos asignados a cada característica.

Histograma	Estadísticos	Dimensión fractal	Gabor	Tamura
5	1	4	6	1
Matriz Co-ocurrencia	DCT	DWT	DFT	
1	3	2	1	

Tabla 43. Pesos de las características del estudio

El valor de los pesos se ha elegido de manera que las cuatro características que menos información proporcionan tienen asignado un peso con valor 1, mientras que las cinco restantes tienen asignados pesos mayores que uno, en orden decreciente, de la que más información aporta a la que menos, en función de los resultados obtenidos en los estudios de inclusión y eliminación de características.

Los resultados para la imagen de ejemplo son los siguientes.

k = 1	k = 3	k = 5	k = 7
0,385	0,376	0,354	0,359

Tabla 44. Tasa de error obtenida para el estudio 1 y diferentes valores de k

Comparando estos resultados con los obtenidos anteriormente se observa una ligera mejora en los valores de la tasa de error, sobre todo para el valor de $k = 5$, para el cual se obtiene el menor valor de la tasa de error de todo el conjunto de pruebas que se han realizado.

6.3.- Análisis de los resultados obtenidos

A continuación se exponen, para cada parte del diseño del sistema que puede modificarse para obtener mejores o peores resultados, las conclusiones obtenidas en base a los estudios realizados.

6.3.1.- Características

El estudio de las características de manera individual lleva a la conclusión de que las características de Gabor y el histograma son las características que individualmente proporcionan mejores resultados. Mientras que en el estudio de la eliminación de las características una por una se observa que, acorde con el resultado obtenido en el estudio de las características por separado, al eliminar las características de Gabor el valor de la tasa de error aumenta considerablemente, no produciéndose este mismo efecto al eliminar el histograma.

De la misma manera, al ver las características de forma individual se aprecia que las que peores resultados ofrecen son los estadísticos y los estadísticos de la matriz de co-ocurrencia. Mientras que si se observa el estudio de la eliminación de las características se

aprecia como es al eliminar la transformada discreta del coseno cuando se obtienen los mejores resultados.

El resto de características no ofrecen un resultado que resalte sobre los demás, ni de manera positiva ni negativa, no variando en gran medida con su inclusión o eliminación. Son características que no suponen un gran coste computacional al sistema, de manera que incluirlas en el sistema no supone un aumento muy grande del tiempo de ejecución.

6.3.2.- Valores de k

En general los mejores resultados se obtienen para $k = 3$, ya que como se ha explicado anteriormente, si el valor de k es demasiado bajo no habrá individuos suficientes para obtener las probabilidades de ocurrencia, y si el valor de k es demasiado grande el número de individuos pertenecientes a cada clase del conjunto de entrenamiento, que no es muy grande en algunos casos, hará que entre los vecinos más cercanos haya obligatoriamente elementos de clases que no son la correcta, ya que el número de vecinos más cercanos puede ser mayor que el propio número de elementos dentro de esa clase.

6.3.3.- Diseño óptimo

El diseño óptimo del sistema tiene la siguiente asignación de pesos y un valor de $k = 5$:

- Histograma: 5
- Estadísticos: 1
- Dimensión fractal: 4
- Características de Gabor: 6
- Características de Tamura: 1
- Matriz de co-ocurrencia: 1
- Transformada discreta del coseno: 3
- Transformada discreta de Wavelet: 2
- Transformada discreta de Fourier: 1

Para este modelo se obtiene una tasa de error del 35,4%.

6.3.4.- Posibles causas de error

Como se puede observar en el Apéndice A, el número imágenes de ambos conjuntos (entrenamiento y test), no se distribuyen de forma heterogénea dentro de las diferentes clases con las que se trabaja, que forman un total de 116. Se puede ver cómo hay clases para las que hay muchísimas más imágenes que para otras, de manera que el algoritmo de aprendizaje no se entrena de manera óptima, lo que puede conducir a una clasificación incorrecta de las clases en las que hay pocos individuos.

En la siguiente figura se muestra la distribución de clases para ambos conjuntos de imágenes.

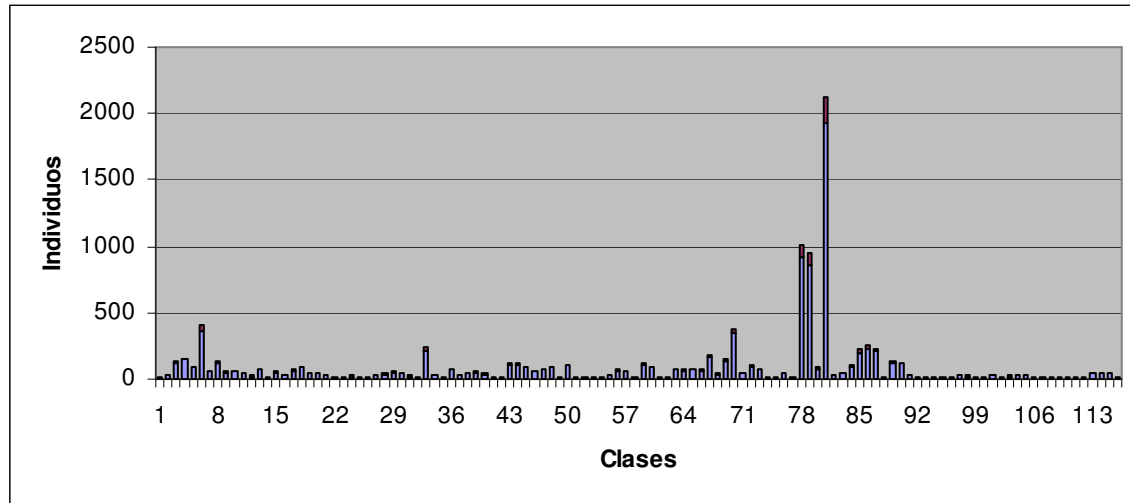


Figura 24. Número de individuos de cada clase

A continuación se muestra en una tabla el número de aciertos y errores cometidos en cada clase para el caso en el que se usan todas las imágenes y $k = 3$:

Clase	Aciertos	Fallos	Total Entrenamiento	Total Test	Tasa de error (%)
1121-127-700-500	0	1	15	1	100
1121-115-710-400	1	3	30	4	75
1121-115-700-400	1	12	125	13	92,31
1121-110-411-700	3	11	144	14	78,57
1121-110-415-700	4	4	87	8	50
1121-110-414-700	29	7	368	36	19,44
1121-120-462-700	4	2	54	6	33,33
1121-200-411-700	5	8	127	13	61,54
1121-120-421-700	2	2	50	4	50
1121-240-421-700	1	5	54	6	83,33
1121-120-442-700	0	4	44	4	100
1121-240-413-700	1	1	21	2	50
1121-120-463-700	3	5	74	8	62,5
1121-12f-467-700	0	2	16	2	100
1121-240-441-700	2	4	49	6	66,67
1121-240-433-700	0	2	30	2	100
1121-490-415-700	3	4	63	7	57,14
1121-4a0-414-700	5	3	83	8	37,5
1121-120-441-700	0	4	46	4	100
1121-240-442-700	3	2	44	5	40
1121-120-433-700	0	2	25	2	100
1121-120-451-700	0	1	14	1	100
1121-120-437-700	0	2	13	2	100
1121-240-437-700	0	3	22	3	100
1121-120-452-700	0	1	12	1	100
1121-240-438-700	0	1	12	1	100
1121-120-413-700	0	2	27	2	100
1121-240-422-700	2	2	35	4	50
1121-230-463-700	0	4	49	4	100
1121-120-422-700	0	5	43	5	100
1121-230-462-700	1	2	20	3	66,67
1121-120-454-700	0	1	11	1	100

6.- Evaluación de resultados

1121-120-800-700	8	13	218	21	61,90
1121-120-951-700	0	3	26	3	100
1121-230-953-700	0	0	10	0	No aplica
1121-230-943-700	1	7	72	8	87,5
1121-120-918-700	1	2	27	3	66,67
1121-120-961-700	1	3	44	4	75
1121-120-962-700	0	6	50	6	100
1121-230-961-700	0	3	37	3	100
1121-120-956-700	0	2	16	2	100
1121-120-950-700	0	2	11	2	100
1121-120-942-700	1	9	106	10	90
1121-120-943-700	7	3	105	10	30
1121-230-942-700	4	5	88	9	55,56
1121-120-911-700	1	4	56	5	80
1121-230-921-700	6	2	74	8	25
1121-120-921-700	5	3	87	8	37,5
1121-230-950-700	0	1	11	1	100
1121-320-941-700	8	1	98	9	11,11
1121-230-915-700	1	1	11	2	50
1121-120-919-700	0	3	19	3	100
1121-230-934-700	0	2	15	2	100
1121-120-933-700	0	0	11	0	No aplica
1121-230-962-700	1	1	29	2	50
1121-120-915-700	4	3	63	7	42,86
1121-120-914-700	1	6	60	7	85,71
1121-230-914-700	0	1	15	1	100
1121-120-922-700	7	4	105	11	36,36
1121-230-922-700	7	2	84	9	22,22
1121-230-911-700	0	2	19	2	100
1121-120-934-700	0	1	20	1	100
1124-310-620-625	5	1	71	6	16,67
1124-310-610-625	4	2	65	6	33,33
1124-410-620-625	5	2	71	7	28,57
1124-410-610-625	6	0	67	6	0
1121-220-230-700	7	9	172	16	56,25
1121-4b0-233-700	0	4	37	4	100
1121-210-230-700	8	6	133	14	42,86
1121-120-200-700	30	3	344	33	9,09
1121-420-213-700	1	3	44	4	75
1121-430-213-700	9	0	93	9	0
1121-220-310-700	6	1	75	7	14,29
1121-4a0-310-700	1	0	17	1	0
1121-490-310-700	0	2	11	2	100
1121-420-212-700	1	4	41	5	80
1121-460-216-700	0	2	15	2	100
1123-110-500-000	64	28	919	92	30,43
1123-211-500-000	73	14	865	87	16,09
1123-121-500-000	4	4	78	8	50
1123-127-500-000	177	18	1927	195	9,23
1121-120-515-700	0	2	27	2	100
1121-120-516-700	0	4	38	4	100
1121-210-310-700	6	3	97	9	33,33
1121-120-310-700	21	0	200	21	0
1121-120-330-700	16	6	230	22	27,27
1121-210-330-700	12	8	213	20	40
1121-210-331-700	1	1	11	2	50
1121-210-320-700	10	2	117	12	16,67
1121-120-320-700	11	1	114	12	8,33
1121-120-311-700	0	3	31	3	100
1121-220-330-700	0	0	10	0	No aplica
1121-229-310-700	0	1	11	1	100
1121-228-310-700	0	0	11	0	No aplica
1121-210-213-700	0	2	12	2	100
1121-120-438-700	0	2	12	2	100
1121-240-434-700	1	2	23	3	66,67

1121-220-213-700	1	1	21	2	50
1121-120-331-700	0	0	10	0	No aplica
1122-220-333-700	0	0	10	0	No aplica
1121-120-517-700	1	2	25	3	66,67
1121-116-917-700	0	1	15	1	100
1121-120-514-700	1	2	22	3	66,67
1121-110-213-700	2	2	32	4	50
1121-127-700-400	0	2	26	2	100
1121-12f-466-700	0	1	13	1	100
1121-120-91a-700	0	1	11	1	100
1121-430-215-700	0	1	11	1	100
1121-230-930-700	0	2	14	2	100
1121-120-434-700	0	2	17	2	100
1121-4a0-918-700	0	2	12	2	100
1121-120-930-700	1	1	15	2	50
1121-490-915-700	1	3	44	4	75
1121-4a0-914-700	2	2	46	4	50
1123-112-500-000	0	5	46	5	100
1121-129-700-400	0	1	14	1	100

Tabla 45. Aciertos y fallos para las diferentes clases

En la tabla anterior vemos como, para las clases que más individuos tienen en el conjunto de entrenamiento, el número de aciertos es mucho mayor que para las clases que tienen pocos individuos en este conjunto. De esta manera podemos ver como las clases con mayor número de individuos en el conjunto de entrenamiento, presentan mejores resultados que las que tienen pocos, tal y como se observa en la siguiente tabla:

Clase	Individuos entrenamiento	Aciertos	Fallos
75	11	0	2
95	12	0	2
34	26	0	3
79	865	73	14
78	919	64	28
81	1927	177	18

Tabla 46. Mejores/peores resultados según individuos entrenamiento

En las siguientes figuras se muestra la tasa de error que se ha obtenido para cada clase de imágenes.

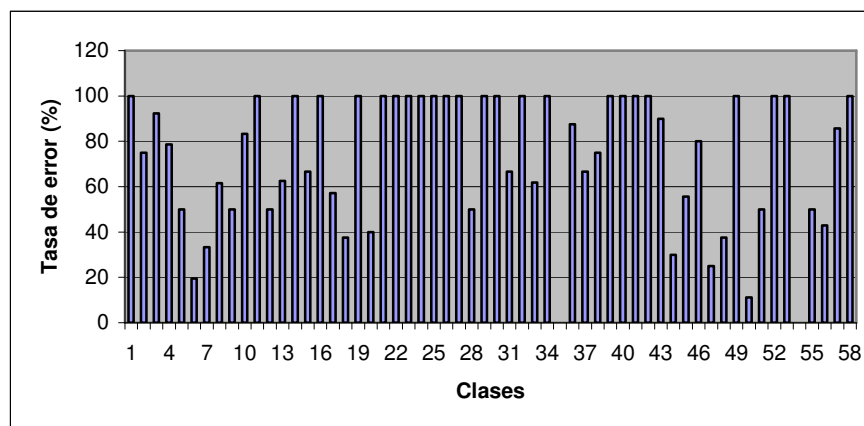


Figura 25. Tasa de error (%) de las clases 1 a 58

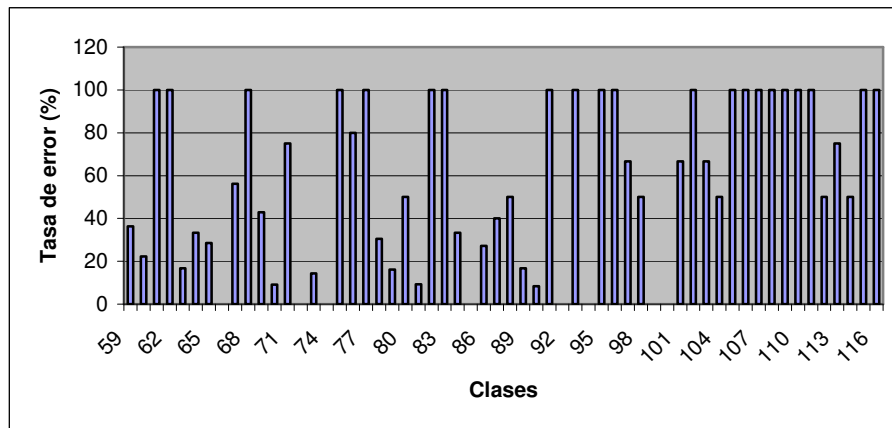


Figura 26. Tasa de error (%) de las clases 59 a 116

Se puede observar, comparando estas dos gráficas con la mostrada en la figura 29 que las clases que menor error presentan se corresponden con las clases que más individuos presentan en el conjunto de imágenes de error. Esto se debe a que el sistema aprende mejor a clasificar los elementos de esas clases debido a que tienen más elementos con los que comparar.

6.3.5.- Comparación con otros sistemas (ImageCLEF 2007)

En este apartado se muestra una comparación de la tasa de error del diseño óptimo del sistema de clasificación de imágenes implementado con los mejores resultados de los grupos que participaron en el congreso ImageCLEF 2007.

En la siguiente figura se observa cómo los resultados asociados con este proyecto (PFC_NMS) alcanzan la novena posición, que si bien no entra dentro del ranking de los cinco mejores, ofrece un resultado bastante aceptable teniendo en cuenta la variabilidad tan pequeña que existe entre los grupos, pudiendo hacer tres distinciones claras entre los cuatro primeros, los cinco siguientes y los dos últimos.

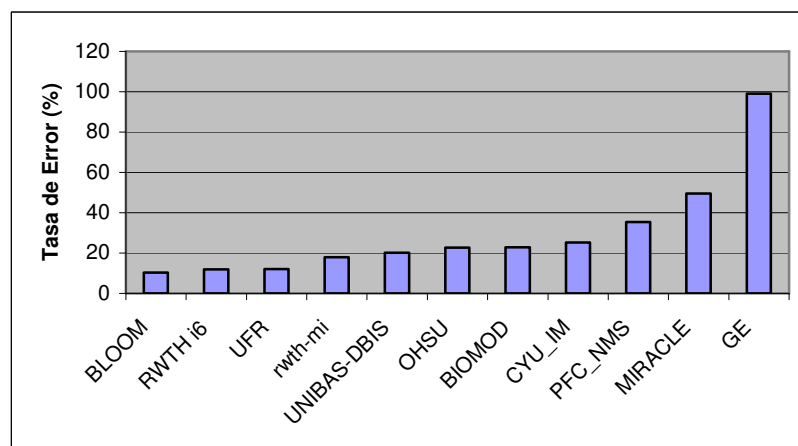


Figura 27. Comparación del sistema con ImageCLEF 2007

Si se recuerda la descripción de los trabajos realizados por los participantes de ImageCLEF 2007 que se hizo en el capítulo del estado del arte, hay que tener en cuenta que en la gráfica anterior sólo se muestran los diez mejores de diecisiete participantes que había, de manera que la novena posición en el ranking es una posición que ofrece unos resultados bastante aceptables. Además, si se compara con el grupo MIRACLE, del departamento de Telemática de la Universidad Carlos III de Madrid, se ve como existe una mejora en la tasa de error, que era en parte lo que se pretendía, al haber sido el estudio de este grupo, en parte, el origen de este proyecyo.

Por último, sobre los grupos que obtienen los mejores resultados cabe destacar, por una parte al grupo BLOOM, que utiliza características SIFT y SUM como clasificador, al igual que el tercer grupo, UFR.

7.- PRESUPUESTO



UNIVERSIDAD CARLOS III DE MADRID
Escuela Politécnica Superior

PRESUPUESTO DE PROYECTO

1.- Autor:

Natalia Muñoz Sarriá

2.- Departamento:

Ingeniería Telemática

3.- Descripción del Proyecto:

- Título: Sistema de clasificación automática de imágenes
- Duración (meses): 14
- Tasa de costes Indirectos: 20%

4.- Presupuesto total del Proyecto (valores en Euros):

37.765,44 Euros

5.- Desglose presupuestario (costes directos)

PERSONAL

Apellidos y nombre	N.I.F.	Categoría	Dedicación (personas) meses ^{a)}	Coste hombre mes	Coste (Euro)
Julio Villena Román	8042900F	Ingeniero Senior	7	4.289,54	30.026,78
Natalia Muñoz Sarriá	2283131J	Ingeniero	0,5	2.694,39	1.347,20
Personas mes 7,5				Total	31.373,98

^{a)} 1 Persona mes = 131,25 horas. Máximo anual de dedicación de 12 personas mes (1575 horas)
Máximo anual para PDI de la Universidad Carlos III de Madrid de 8,8 personas mes (1.155 horas)

EQUIPOS

Descripción	Coste (Euro)	% Uso dedicado proyecto	Dedicación (meses)	Periodo de depreciación	Coste imputable ^{d)}
Ordenador portátil	800,00	100	7	60	93,33
Disco de backup	200,00	100	7	60	23,33
				60	0,00
Total					116,67

^{d)} Fórmula de cálculo de la Amortización:

$$\frac{A}{B} \times C \times D$$

A = nº de meses desde la fecha de facturación en que el equipo es utilizado
B = periodo de depreciación (60 meses)
C = coste del equipo (sin IVA)
D = % del uso que se dedica al proyecto (habitualmente 100%)

SUBCONTRATACIÓN DE TAREAS

Descripción	Empresa	Coste imputable
Total		0,00

OTROS COSTES DIRECTOS DEL PROYECTO^{e)}

Descripción	Empresa	Costes imputable
Total		0,00

^{e)} Este capítulo de gastos incluye todos los gastos no contemplados en los conceptos anteriores, por ejemplo: fungible, viajes y dietas, otros,...

6.- Resumen de costes

Presupuesto Costes Totales	Presupuesto Costes Totales
Personal	31.374
Amortización	117
Subcontratación de tareas	0
Costes de funcionamiento	0
Costes Indirectos	6.275
Total	37.765

8.- CONCLUSIONES Y TRABAJOS FUTUROS

8.1.- Introducción

En el presente capítulo se presentan tanto las conclusiones finales una vez diseñado, implementado y evaluado el funcionamiento del sistema, como los posibles trabajos con los que se podría continuar trabajando en búsqueda de un sistema de clasificación de imágenes basado en contenido óptimo.

8.2.- Conclusiones obtenidas de la realización del proyecto

Tal y como se explicó desde el primer capítulo, el objetivo de este proyecto fue la implementación de un sistema de clasificación automática de imágenes médicas, codificadas mediante un código denominado IRMA, de estructura mono-jerárquica, completo e inequívoco.

En primer lugar, mediante el desarrollo del proyecto se ha realizado un estudio de las características que pueden obtenerse de las imágenes con las que se ha trabajado, llegando a identificar, de entre las que se estudiaron en el capítulo del estado del arte, las que más y las que menos información aportan a la hora de clasificar las imágenes.

En segundo lugar se estudió un algoritmo de clasificación, en este caso concreto el algoritmo k-Nearest Neighbour (vecino más cercano), y sus posibles implementaciones de manera que se pudieran obtener los mejores resultados posibles (se han estudiado la diferencia en los resultados finales al utilizar distintos números de vecinos más cercanos, y los motivos de la variación en estos resultados para los diferentes valores).

Para la evaluación del sistema se ha empleado una métrica sencilla como es la tasa de error, es decir, el número de errores sobre el total de casos estudiados.

Los estudios realizados han permitido observar los resultados obtenidos para diferentes combinaciones de características (estudio de las características por separado, de todas las características, eliminación de las características una por una y la ponderación de las mismas con pesos distintos de 0 y 1) en la etapa de extracción de las mismas, y diferentes valores del número de vecinos más cercanos en la etapa de clasificación ($k = 1, 3, 5$ y 7). De esta manera se ha llegado al diseño óptimo con el cual se ha obtenido el mejor resultado de la tasa de error.

El estudio de las características de manera individual lleva a la conclusión de que las características de Gabor y el histograma son las características que individualmente proporcionan mejores resultados. Mientras que en el estudio de la eliminación de las características una por una se observa que, acorde con el resultado obtenido en el estudio de

las características por separado, al eliminar las características de Gabor el valor de la tasa de error aumenta considerablemente, no produciéndose este mismo efecto al eliminar el histograma.

De la misma manera, al ver las características de forma individual se aprecia que las que peores resultados ofrecen son los estadísticos y los estadísticos de la matriz de co-ocurrencia. Mientras que si se observa el estudio de la eliminación de las características se aprecia como es al eliminar la transformada discreta del coseno cuando se obtienen los mejores resultados.

El resto de características no ofrecen un resultado que resalte sobre los demás, ni de manera positiva ni negativa, no variando en gran medida con su inclusión o eliminación. Son características que no suponen un gran coste computacional al sistema, de manera que incluirlas en el sistema no supone un aumento muy grande del tiempo de ejecución.

En general los mejores resultados se obtienen para $k = 3$, ya que como se ha explicado anteriormente, si el valor de k es demasiado bajo no habrá individuos suficientes para obtener las probabilidades de ocurrencia; y si el valor de k es demasiado grande el número de individuos pertenecientes a cada clase del conjunto de entrenamiento, que no es muy grande en algunos casos, hará que entre los vecinos más cercanos haya obligatoriamente elementos de clases que no son la correcta, ya que el número de vecinos más cercanos puede ser mayor que el propio número de elementos dentro de esa clase.

El diseño óptimo del sistema se ha encontrado en el “Estudio 1” con la siguiente asignación de pesos y un valor de $k = 5$:

- Histograma: 5
- Estadísticos: 1
- Dimensión fractal: 4
- Características de Gabor: 6
- Características de Tamura: 1
- Matriz de co-ocurrencia: 1
- Transformada discreta del coseno: 3
- Transformada discreta de Wavelet: 2
- Transformada discreta de Fourier: 1

Para este modelo se obtiene una tasa de error del 35,4%.

Por último, se ha realizado una comparativa con los grupos que participaron en la tarea de ImageCLEF 2007, en los que se observa como los resultados asociados con este proyecto alcanzan la novena posición, que si bien no entra dentro del ranking de los cinco primeros,

ofrece un resultado bastante aceptable teniendo en cuenta el tiempo y esfuerzo limitados de un proyecto fin de carrera, y la técnica sencilla empleada en este sistema.

8.3.- Trabajos futuros

En general, en la elaboración de un proyecto no se alcanza un resultado ideal ya que no es posible abarcar el conjunto completo de pruebas, por varios motivos, como puede ser no tener los recursos suficientes o el tiempo suficiente (en algunos casos este tiempo podría incluso tender a infinito).

Por estos motivos siempre puede haber posibles ampliaciones del mismo que, en muchos casos (aunque no necesariamente en todos), acercarán los resultados a los óptimos. En este caso se han identificado las siguientes posibles ampliaciones.

A) Mejoras en la etapa de extracción de características

Tal y como se explicó en el capítulo de estado del arte, existe una enorme cantidad de características distintas que se pueden obtener de las imágenes. Además dentro de estas características hay muchas en las que existen parámetros variables que permiten variar los resultados.

- Introducción de más características: las siguientes características se describieron en el capítulo del estado del arte y posteriormente no se incluyeron en el diseño del sistema:
 - Características de Laws.
 - Características invariantes.
 - Histograma de características invariantes.
 - Estadísticos de características invariantes.
 - Características invariantes de Fourier-Mellin.
 - Dilatación de la imagen previa a la extracción de características.
 - Erosión de la imagen previa a la extracción de características.
 - Filtro de apertura de la imagen previa a la extracción de características.
 - Filtro de cierre de la imagen previa a la extracción de características.
 - Transformada de Karhunen-Loève.
 - Transformada de Hough.
- Estudio de diferentes valores de los parámetros variables de las características (en concreto, de las características añadidas en el diseño del sistema, estos parámetros se describieron en el capítulo de diseño del sistema, algunos ejemplo pueden ser, el número de niveles de gris del histograma, el número de orientaciones y escalas de las características de Gabo, etc)

La inclusión de nuevas características y nuevos valores de los parámetros variables permiten tener una cantidad de información que comparar mayor, lo que puede ayudar a alcanzar mejores resultados a la hora de la clasificación de las imágenes.

B) Mejoras en la extracción de características locales

En el sistema implementado se han dividido las imágenes (previamente redimensionadas a 256x256 píxeles) en bloques de 128x128 píxeles (por tanto en cuatro bloques).

Para hacer una comparativa mejor de las imágenes y sus diferentes detalles puede aumentarse este número de bloques mediante la disminución del tamaño de los mismos. De esta manera quizá se podría, tal y como se explicó en el capítulo de desarrollo y pruebas, disminuir el número de errores cometidos en el eje de dirección y en el eje anatómico.

C) Aumento del número de combinaciones de características estudiadas

También se vió en el capítulo de desarrollo y pruebas que el número de combinaciones de características es prácticamente inabarcable, ya que, en caso de que se usen pesos 0 y 1, se tendrán 2^N posibles combinaciones, y si se usan pesos que varíen de 0 a 9 se tendrán 10^N posibles combinaciones, siendo N es número de características utilizadas en la comparación.

D) Mejoras en el algoritmo de clasificación

- Además de la distancia euclídea existen otro tipo de distancias, como puede ser la distancia de Mahalanobis, que podría proporcionar mejores resultados ya que tiene en cuenta parámetros estadísticos como la media y la varianza de las características, a cambio de un aumento en el coste computacional de su cálculo.
- Implementación de un algoritmo de clasificación distinto del k-NN, como podría ser una máquina de vectores soporte (SVM - Support Vector Machine), de manera que podrían obtenerse mejores resultados a cambio de un aumento del coste computacional en la fase de entrenamiento (aunque no en la fase de test).
- Modificar las pautas de decisión a la hora de determinar qué código se le asigna de entre los posibles (se tendrán k códigos distintos como mucho, a no ser que alguno se repita dentro de los k vecinos más cercanos, en cuyo caso el número de códigos posibles será menor que k). Para asignar un código a una imagen se siguen los siguientes pasos:

8.- Conclusiones y trabajos futuros

1.- Obtención de los k vecinos más cercanos a la imagen de test y su correspondiente código IRMA. Una vez determinados cuáles son los k vecinos más cercanos de la imagen de test, se almacena el código IRMA de cada uno de ellos.

k CODIGOS IRMA DE LOS k VECINOS MAS CERCANOS			
TECNICO	DIRECCIONAL	ANATOMICO	BIOLOGICO
1121	120	200	700
1121	120	800	700
1121	420	213	700
1121	120	800	700
1121	120	200	700

Tabla 47. Ejemplo de k códigos IRMA de los k vecinos más cercanos a una imagen de test

2.- Determinar la probabilidad del código de cada eje. Es la probabilidad de ocurrencia de ese código dentro de los k vecinos más cercanos, de manera que sigue la siguiente expresión:

$$probabilidad_codigo = \frac{n^{\circ} \text{ocurrencias}}{k}$$

No se usa la probabilidad del código completo, sino del subcódigo que pertenece a cada uno de los cuatro ejes que conforman el código IRMA de cada imagen.

TECNICO		DIRECCIONAL		ANATOMICO		BIOLOGICO	
Código	Probab.	Código	Probab.	Código	Probab.	Código	Probab.
1121	5/5 = 1	120	4/5	200	2/5	700	5/5 = 1
		420	1/5	800	2/5		
				213	1/5		

Tabla 48. Ejemplo de k códigos IRMA con las probabilidades por eje

3.- Encontrar los N posibles códigos IRMA que se pueden formar mediante combinación, con su correspondiente probabilidad. En este paso, se despliegan todos los posibles códigos estableciendo todas las posibles combinaciones de los ejes de los k vecinos y se calcula la probabilidad de obtención de cada uno.

N CODIGOS IRMA DE LOS k VECINOS MAS CERCANOS				PROBABILIDAD
TECNICO	DIRECCIONAL	ANATOMICO	BIOLOGICO	
1121	120	200	700	0.32
1121	120	800	700	0.32
1121	120	213	700	0.16
1121	420	200	700	0.08
1121	420	800	700	0.08
1121	420	213	700	0.04

Tabla 49. Ejemplo de N posibles códigos al combinar los k códigos de los vecinos más cercanos

4.- Establecer cuáles son códigos IRMA válidos y sus correspondientes probabilidades.

En el paso anterior, es muy frecuente que se formen códigos que no sean válidos, ya que no siempre es posible encontrar cualquier combinación entre dos ejes, más si cabe entre los cuatro. En este paso, se deben eliminar los códigos que sean inválidos, y repartir de forma equitativa las probabilidades de los códigos eliminados entre los códigos que finalmente si son válidos.

Llegados a este punto el resultado será el código que mayor probabilidad de ocurrencia presente. En caso de empate habrá que definir una regla que nos permitiera elegir uno u otro.

N CODIGOS IRMA DE LOS k VECINOS MAS CERCANOS				PROBABILIDAD
TECNICO	DIRECCIONAL	ANATOMICO	BIOLOGICO	
1121	120	200	700	42.667
1121	120	800	700	42.667
1121	420	213	700	14.667

Tabla 50. Ejemplo de los posibles códigos validos de una imagen de test

E) Mejoras en el método de evaluación de resultados

En lugar de considerar acierto el código completo se puede evaluar por eje de manera independiente.

F) Mejora de la base de datos de imágenes

Se ha visto que la heterogeneidad de la base de datos utilizada puede llevar a cometer errores frecuentes en las clases en las que hay pocos elementos muestrales. La elección de una base de datos más heterogénea puede reducir este número de errores.

APÉNDICES

Apéndice A. Clases IRMA

Clase	Codigo	Entrenamiento	Test	Clase	Codigo	Entrenamiento	Test
1	1121-127-700-500	15	1	59	1121-120-922-700	105	11
2	1121-115-710-400	30	4	60	1121-230-922-700	84	9
3	1121-115-700-400	125	13	61	1121-230-911-700	19	2
4	1121-110-411-700	144	14	62	1121-120-934-700	20	1
5	1121-110-415-700	87	8	63	1124-310-620-625	71	6
6	1121-110-414-700	368	36	64	1124-310-610-625	65	6
7	1121-120-462-700	54	6	65	1124-410-620-625	71	7
8	1121-200-411-700	127	13	66	1124-410-610-625	67	6
9	1121-120-421-700	50	4	67	1121-220-230-700	172	16
10	1121-240-421-700	54	6	68	1121-4b0-233-700	37	4
11	1121-120-442-700	44	4	69	1121-210-230-700	133	14
12	1121-240-413-700	21	2	70	1121-120-200-700	344	33
13	1121-120-463-700	74	8	71	1121-420-213-700	44	4
14	1121-12f-467-700	16	2	72	1121-430-213-700	93	9
15	1121-240-441-700	49	6	73	1121-220-310-700	75	7
16	1121-240-433-700	30	2	74	1121-4a0-310-700	17	1
17	1121-490-415-700	63	7	75	1121-490-310-700	11	2
18	1121-4a0-414-700	83	8	76	1121-420-212-700	41	5
19	1121-120-441-700	46	4	77	1121-460-216-700	15	2
20	1121-240-442-700	44	5	78	1123-110-500-000	919	92
21	1121-120-433-700	25	2	79	1123-211-500-000	865	87
22	1121-120-451-700	14	1	80	1123-121-500-000	78	8
23	1121-120-437-700	13	2	81	1123-127-500-000	1927	195
24	1121-240-437-700	22	3	82	1121-120-515-700	27	2
25	1121-120-452-700	12	1	83	1121-120-516-700	38	4
26	1121-240-438-700	12	1	84	1121-210-310-700	97	9
27	1121-120-413-700	27	2	85	1121-120-310-700	200	21
28	1121-240-422-700	35	4	86	1121-120-330-700	230	22
29	1121-230-463-700	49	4	87	1121-210-330-700	213	20
30	1121-120-422-700	43	5	88	1121-210-331-700	11	2
31	1121-230-462-700	20	3	89	1121-210-320-700	117	12
32	1121-120-454-700	11	1	90	1121-120-320-700	114	12
33	1121-120-800-700	218	21	91	1121-120-311-700	31	3
34	1121-120-951-700	26	3	92	1121-220-330-700	10	0
35	1121-230-953-700	10	0	93	1121-229-310-700	11	1
36	1121-230-943-700	72	8	94	1121-228-310-700	11	0
37	1121-120-918-700	27	3	95	1121-210-213-700	12	2
38	1121-120-961-700	44	4	96	1121-120-438-700	12	2
39	1121-120-962-700	50	6	97	1121-240-434-700	23	3
40	1121-230-961-700	37	3	98	1121-220-213-700	21	2
41	1121-120-956-700	16	2	99	1121-120-331-700	10	0
42	1121-120-950-700	11	2	100	1122-220-333-700	10	0
43	1121-120-942-700	106	10	101	1121-120-517-700	25	3
44	1121-120-943-700	105	10	102	1121-116-917-700	15	1
45	1121-230-942-700	88	9	103	1121-120-514-700	22	3
46	1121-120-911-700	56	5	104	1121-110-213-700	32	4
47	1121-230-921-700	74	8	105	1121-127-700-400	26	2
48	1121-120-921-700	87	8	106	1121-12f-466-700	13	1
49	1121-230-950-700	11	1	107	1121-120-91 ^a -700	11	1
50	1121-320-941-700	98	9	108	1121-430-215-700	11	1
51	1121-230-915-700	11	2	109	1121-230-930-700	14	2
52	1121-120-919-700	19	3	110	1121-120-434-700	17	2
53	1121-230-934-700	15	2	111	1121-4a0-918-700	12	2
54	1121-120-933-700	11	0	112	1121-120-930-700	15	2
55	1121-230-962-700	29	2	113	1121-490-915-700	44	4
56	1121-120-915-700	63	7	114	1121-4a0-914-700	46	4
57	1121-120-914-700	60	7	115	1123-112-500-000	46	5
58	1121-230-914-700	15	1	116	1121-129-700-400	14	1

Apéndice B. Códigos IRMA

A Technical code for imaging modality

0 unspecified

1 x-ray

10 unspecified

11 plain radiography

110 unspecified

111 digital

1110 unspecified

1111 overview image

1112 tomography

1113 high beam energy

1114 low beam energy

1115 far view projection

1116 1:1 projection

1117 dual energy

112 analog

1120 unspecified

1121 overview image

1122 tomography

1123 high beam energy

1124 low beam energy

1125 far view projection

1126 1:1 projection

113 x-ray stereometry

114 x-ray stereography

115 tomosynthesis

116 xeroradiography

12 fluoroscopy

120 unspecified

121 digital

122 analog

13 angiography

130 unspecified

131 digital

1310 unspecified

1311 subtraction

1312 rotation

1313 pathfinder

132 analog

1320 unspecified

1321 photo subst

14 computed tomography (CT)

140 unspecified

141 conventional CT

1410 unspecified

1411 low dose

1412 high resolution

1413 bolus tracking

142 quantitative CT

1420 unspecified

1421 dual energy

15 DEXA (quantitative radiography)

150 unspecified

151 digital

1510 unspecified

1511 dual energy

16 radiotherapy

160 unspecified

161 digital

162 analog

2 sonography

20 unspecified

21 A-mode

210 unspecified

221 3.5 MHz

212 5.0 MHz

213 7.5 MHz

214 10 MHz

215 variable frequency

22 B-mode

220 unspecified

221 sector scan

2210 unspecified

2211 3.5 MHz

2212 5.0 MHz

2213 7.5 MHz

2214 10 MHz

2215 variable freq

222 linear scan

2220 unspecified

2221 3.5 MHz

2222 5.0 MHz

2223 7.5 MHz

2224 10 MHz

2225 variable freq

23 M-mode

230 unspecified	3110 unspecified
231 3.5 MHz	3111 endo coil
232 5.0 MHz	3112 body
233 7.5 MHz	3113 body array
234 10 MHz	3114 head coil
235 variable freq	3115 neck coil
24 compound mode	3116 spine coil
240 unspecified	3117 loop coil
241 3.5 MHz	3118 shoulder coil
242 5.0 MHz	3119 knee coil
243 7.5 MHz	311a hand / small FOV
244 10 MHz	311b mamma coil
245 variable freq	311c flexible coil, large
25 SieScape mode	311d flexible coil, small
250 unspecified	311e endorectal coil
251 3.5 MHz	311f wrap around coil
252 5.0 MHz	312 T2-weighted
253 7.5 MHz	3120 unspecified
254 10 MHz	3121 endo coil
255 variable freq	3122 body
26 duplex mode	3123 body array
260 unspecified	3124 head coil
261 3.5 MHz	3125 neck coil
262 5.0 MHz	3126 spine coil
263 7.5 MHz	3127 loop coil
264 10 MHz	3128 shoulder coil
265 variable freq	3129 knee coil
27 Doppler mode	312a hand / small FOV
270 unspecified	312b mamma coil
271 3.5 MHz	312c flexible coil, large
272 5.0 MHz	312d flexible coil, small
273 7.5 MHz	312e endorectal coil
274 10 MHz	312f wrap around coil
275 variable freq	313 proton-weighted
28 3D imaging	3130 unspecified
280 unspecified	3131 endo coil
281 3.5 MHz	3132 body
282 5.0 MHz	3133 body array
283 7.5 MHz	3134 head coil
284 10 MHz	3135 neck coil
285 variable frequency	3136 spine coil
3 magnetic resonance measurements	3137 loop coil
30 unspecified	3138 shoulder coil
31 magnetic resonance imaging (MRI)	3139 knee coil
310 unspecified	313a hand / small FOV
311 T1-weighted	313b mamma coil

313c flexible coil, large	3167 loop coil
313d flexible coil, small	3168 shoulder coil
313e endorectal coil	3169 knee coil
313f wrap around coil	316a hand / small FOV
314 fat suppression	316b mamma coil
3140 unspecified	316c flexible coil, large
3141 endo coil	316d flexible coil, small
3142 body	316e endorectal coil
3143 body array	316f wrap around coil
3144 head coil	317 FLARE
3145 neck coil	3170 unspecified
3146 spine coil	3171 endo coil
3147 loop coil	3172 body
3148 shoulder coil	3173 body array
3149 knee coil	3174 head coil
314a hand / small FOV	3175 neck coil
314b mamma coil	3176 spine coil
314c flexible coil, large	3177 loop coil
314d flexible coil, small	3178 shoulder coil
314e endorectal coil	3179 knee coil
314f wrap around coil	317a hand / small FOV
315 Cine-Lolo	317b mamma coil
3150 unspecified	317c flexible coil, large
3151 endo coil	317d flexible coil, small
3152 body	317e endorectal coil
3153 body array	317f wrap around coil
3154 head coil	318 3D-FFE high resolution
3155 neck coil	3180 unspecified
3156 spine coil	3181 endo coil
3157 loop coil	3182 body
3158 shoulder coil	3183 body array
3159 knee coil	3184 head coil
315a hand / small FOV	3185 neck coil
315b mamma coil	3186 spine coil
315c flexible coil, large	3187 loop coil
315d flexible coil, small	3188 shoulder coil
315e endorectal coil	3189 knee coil
315f wrap around coil	318a hand / small FOV
316 inversion recovery	318b mamma coil
3160 unspecified	318c flexible coil, large
3161 endo coil	318d flexible coil, small
3162 body	318e endorectal coil
3163 body array	318f wrap around coil
3164 head coil	32 spectroscopy
3165 neck coil	4 nuclear medicine
3166 spine coil	40 unspecified

41 scintillation probe	11 posteroanterior (PA)
42 gamma camera	110 unspecified
43 single photon emission computed tomography (SPECT)	111 inspiration
44 positron emission tomography (PET)	112 expiration
5 optical imaging	113 valsalva
50 unspecified	114 phonation
51 endoscopy	115 upright
52 microscopy	116 sitting
53 photography	117 supine
54 funduscopy	118 prone
55 laser surface scan	119 lateral decubitus
56 thermography	11a flexion, left
6 biophysical procedures	11b flexion, right
60 unspecified	11c swallow
61 electric	11d micturition
62 magnetic	11e bending
63 probes	12 anteroposterior (AP, coronal)
7 others	120 unspecified
70 unspecified	121 inspiration
71 workstation	122 expiration
72 hardcopy	123 valsalva
73 secondary capture	124 phonation
74 documentation	125 upright
740 unspecified	126 sitting
741 value	127 supine
742 curve local	128 prone
743 curve temporal	129 lateral decubitus
744 area complex (3D)	12a flexion, left
745 picture	12b flexion, right
746 local series	12c swallow
747 temporal series	12d micturition
748 complex series (4D)	12e bending
749 cine documentation	2 sagittal
74a fluoroscopy	20 unspecified
74b volume (data set)	21 lateral, right-left
74c volume temporal (data set)	210 unspecified
8 secondary digitalization	211 inspiration
80 unspecified	212 expiration
81 laser sampling	213 valsalva
82 LCD sampling	214 Hitzenberg maneuver
83 video digitizing	215 upright
B Directional code for imaging orientation	216 supine
0 unspecified	217 prone
1 coronal	218 inclination
10 unspecified	219 reclination
	21a swallow

21b defaecation	212 eye area
21c micturition	213 nose area
22 lateral, left-right	214 maxilla
220 unspecified	215 mandible
221 inspiration	216 temporo mandibular area
222 expiration	22 cranial base
223 valsalva	220 unspecified
224 Hitzenberg maneuver	221 petrous bone
225 upright	222 sella
226 supine	23 neuro cranium
227 prone	230 unspecified
228 inclination	231 frontal area
229 reclination	232 parietal area
22a swallow	233 occipital area
22b defaecation	3 spine
22c micturition	30 unspecified
23 mediolateral	31 cervical spine
24 lateromedial	310 unspecified
3 axial	311 dens
30 unspecified	312 axis
31 craniocaudal	313 upper cervical spine
32 caudocranial (transversal, axial)	314 lower cervical spine
4 other orientation	32 thoracic spine
40 unspecified	320 unspecified
41 oblique	321 cervico-thoracic conjunctio
42 occipitofrontal	322 upper thoracic spine
43 occipitomental	323 middle thoracic spine
44 bregmaticooccipital	324 lower thoracic spine
45 submentobregmatical	33 lumbar spine
46 submentobregmaticofrontal	330 unspecified
47 bregmaticooral	331 thoraco-lumbar conjunctio
48 bregmaticosubmental	332 upper lumbar spine
49 right anterior oblique (RAO)	333 lower lumbar spine
4a left anterior oblique (LAO)	34 sacral bone
5 RSA / ap and lateral (simultaneously)	340 unspecified
C Anatomical code for body region examined	341 lumbo-sacral conjunction
0 unspecified	35 coccygeal bone
1 whole body	4 upper extremity / arm
10 unspecified	40 unspecified
11 torso	41 hand
12 extremities	410 unspecified
2 cranium	411 finger
20 unspecified	412 middle hand
21 facial cranium	413 carpal bones
210 unspecified	42 radio carpal joint
211 forehead	43 forearm

430 unspecified	722 peri navel region
431 distal forearm	723 middle left quadrant
432 proximal forearm	73 lower abdomen
44 elbow	730 unspecified
45 upper arm	731 lower right quadrant
450 unspecified	732 lower middle quadrant
451 distal upper arm	733 lower left quadrant
452 proximal upper arm	8 pelvis
46 shoulder	80 unspecified
460 unspecified	81 sacral bone
461 scapula	810 unspecified
462 humero-scapular joint	811 iliosacral-junction
463 acromio-scapula joint	82 iliac bone
5 chest	83 pubic bone
50 unspecified	830 unspecified
51 bones	831 symphysis
510 unspecified	84 small pelvis
511 clavicle	840 unspecified
512 sternoclavicle region	841 sacral bone
513 sternum	842 ischial bone
514 upper ribs	9 lower extremity / leg
515 lower ribs	90 unspecified
52 lung	91 foot
520 unspecified	910 unspecified
521 upper lobe	911 toe
522 middle lobe	912 middle foot
523 lower lobe	913 tarsal bones
53 hilum	92 ankle joint
54 mediastinum	93 lower leg
540 unspecified	930 unspecified
541 anterior mediastinum	931 distal lower leg
542 middle mediastinum	932 proximal lower leg
543 posterior mediastinum	94 knee
55 heart	940 unspecified
56 diaphragm	941 patella
6 breast (mamma)	95 upper leg
7 abdomen	950 unspecified
70 unspecified	951 distal upper leg
71 upper abdomen	952 proximal upper leg
710 unspecified	96 hip
711 upper right quadrant	D Biological code for system examined
712 upper middle quadrant	0 unspecified
713 upper left quadrant	1 cerebrospinal system
72 middle abdomen	10 unspecified
720 unspecified	11 central nervous system
721 middle right abdomen	110 unspecified

111 metencephalon, cerebrum	225 botalli ligament & ductus arteriosus
112 mesencephalon, diencephalon	226 descending aorta
113 neurohypophysis, posterior pituitary gland	227 bronchial trunk
114 cerebellum	228 abdominal aorta
115 pons	229 celiac trunk
116 medulla oblongata	22a common iliac artery
117 spinal cord, cervical	22b other arteries
118 spinal cord, thoracic	23 veins
119 spinal cord, lumbar	230 unspecified
11a spinal cord, sacral	231 anonymus vein
12 sense organs	232 brachiocephalic vein
120 unspecified	233 superior vena cava
121 vestibular sense	234 inferior vena cava
122 hearing sense	235 common iliac vein
123 gustatory sense	236 other veins
124 tactile sense	3 respiratory system
125 vision sense	30 unspecified
126 olfactory sense	31 nose
13 vegetative nervous system	32 sinuses
130 unspecified	33 nasopharynx & epipharynx
131 sympathetic chain	34 mesopharynx
132 glomus jugulare	35 hypopharynx
133 glomus caroticum	36 larynx
134 ganglion stellatum	37 trachea
135 ganglion coeliacum	38 main bronchus
136 organ of Zuckerkandl	39 bronchi
137 adrenal medulla	3a lung
2 cardiovascular system	3b pleura
20 unspecified	3c ribs
21 heart	3d diaphragma
210 unspecified	3e intercostal muscles
211 pericardium	4 gastrointestinal system
212 atrium	40 unspecified
213 mitral valve	41 oropharynx
214 tricuspidal valve	410 unspecified
215 ventricle	411 oral cavity
216 aortic valve	412 epipharynx
217 pulmonary valve	413 mesopharynx
22 arteries	414 hypopharynx
220 unspecified	42 esophagus
221 ascending aorta	420 unspecified
222 aortic arche	421 upper esophagus sphincter
223 brachiocephalic trunk	422 upper third
224 common carotic artery	423 middle third

424 lower third	510 unspecified
425 lower esophagus	511 parenchyma
sphincter	512 renal pelvis
43 stomach	52 ureter
430 unspecified	53 urinary bladder
431 cardia	54 urethra
432 fundus	6 reproductive system
433 corpus	60 unspecified
434 antrum	61 male system
435 pylorus	610 unspecified
44 small intestine	611 testes
440 unspecified	612 deferent duct
441 duodenum	613 prostate gland
442 jejunum	614 penis
443 ileum	62 female system
444 terminal ileum	620 unspecified
445 ileoceocal valve	621 ovaries
45 large intestine	622 tuba uterina
450 unspecified	623 uterus
451 ascending colon	624 vagina
452 transverse colon	625 breast
453 descending colon	7 musculoskeletal system
454 sigmoid colon	8 endocrinic system
455 rectum	80 unspecified
46 appendix	81 thyroid gland
47 anus	82 parathyroid gland
48 liver	83 adrenal gland
480 unspecified	84 pituitary gland
481 parenchyma	85 hypothalamus
482 bile ducts	86 ovaries
483 portal vein	87 testes
484 arteries	88 pancreas
485 veins	9 immunic system
49 biliary system	90 unspecified
490 unspecified	91 tonsils
491 bile ducts	92 cervical lymph node
492 gallbladder	93 axillary lymph node
4a salivary glands	94 mediastinal lymph node
4a0 unspecified	95 thymus
4a1 parotic gland	96 spleen
4a2 sublingual gland	97 retroperitoneal lymph node
4a3 mandibular gland	98 intraperitoneal lymph node
4a4 pancreas	980 unspecified
5 uropoietic system	981 mesenterial lymph node
50 unspecified	99 iliac lymph node
51 kidney	9a inguinal lymph node

BIBLIOGRAFÍA Y REFERENCIAS

- [1] Alonso Martín, A.B.; Cid Sueiro, J.; Alba Castro, J.L.; Martín Rodríguez, F.; Navia Vázquez, A. (2007). IMAGine: Cursos Interactivos de Tratamiento Digital de Imagen. Transformadas. Versión 2.0. Consultado en Abril de 2010 en http://www.tsc.uc3m.es/imagine/Curso_ProcesadoBasico/index.html.
- [2] Ángeles Larrieta, María Isabel y Santillán Gómez, Angélica María (2003). Minería de datos: concepto, características, estructura y aplicaciones. Consultado en Septiembre de 2009 en <http://www.ejournal.unam.mx/rca/190/RCA19007.pdf>
- [3] Arrúa, Luciana; Meza Fernández, Eduardo (2003). Inteligencia Artificial: Sistemas expertos y Redes neuronales. Consultado en Marzo de 2009 en <http://exa.unne.edu.ar/depar/areas/informatica/SistemasOperativos/InteligenciaArtificial.PDF>
- [4] B.S. Manjunath; Philippe Salembier; Thomas Sikora. (2002) Introduction to MPEG-7: Multimedia Content Description Interface. Consultado en Diciembre 2009 en <http://books.google.es/books?id=CmSPGXF1yB4C&lpg=PP19&ots=p1W4hKY4R3&dq=Introduction%20to%20MPEG-7%3A%20Multimedia%20Content%20Description%20Interface&lr&pg=PP36#v=onepage&q&f=false>
- [5] B. S. Reddy, B. Chatterji (1996). An FFT-Based Technique for Translation, Rotation and Scale invariant Image Registration. IEEE Transactions on Image Processing. Consultado en Diciembre 2009.
- [6] Barnsley, M. (1988). Fractals everywhere. Consultado en Enero 2010 en <http://books.google.es/books?id=oh7NoePgmOIC&lpg=PR11&ots=PgSgMQQyh-&dq=Fractals%20everywhere&lr&pg=PA3#v=onepage&q&f=false>
- [7] Benoît Mandelbrot (1997). La Geometría Fractal de la Naturaleza. Consultado en Enero 2010 en <http://en.scientificcommons.org/6987559>
- [8] Bernal López, Edwin Andrés (2005). Inteligencia Artificial aplicada al diagnóstico médico. Consultado en Septiembre de 2009 en es.geocities.com/edwin99109/final_arte1.pdf
- [9] Botella, Ana (2003). Redes Neuronales. Consultado en Septiembre de 2009 en <http://www.monografias.com/trabajos12/redneuro/redneuro.shtml#intro>

- [10] Caicedo Rueda, Juan Carlos (2006). Prototipo de Sistema para el almacenamiento y recuperación por contenido de imágenes médicas de histopatología. Consultado en Septiembre de 2009 en <http://dis.unal.edu.co/profesores/ypinzon/2013326-206/docs/Tesis1Caicedo.pdf>
- [11] Caicedo Rueda, Juan Carlos (2006). Extracción de Características para recuperación de Imágenes por Contenido. Consultado en Abril 2010 en <http://disi.unal.edu.co/~ypinzon/2013326-206/docs/Articulo1Caicedo.pdf>
- [12] Cheng PC, Chien BC, Ke HR, Yang WP (2005). NCTU_DBLAB@ImageCLEF2005: Automatic Annotation Task. Consultado en Octubre de 2009.
- [13] Cheng PC, Yang WP (2007). CYU_IM@ImageCLEF 2007: Medical image annotation task. Working Notes of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.
- [14] Corzo, Yuliana (2001). La lógica difusa. Consultado en Septiembre de 2009 en http://www.galeon.com/casanchi/casanchi_2001/difusa01.htm
- [15] Coto, Ernesto (2005). Métodos de segmentación de imágenes médicas. Consultado en Septiembre de 2009 en <http://lcg.ciens.ucv.ve/~ernesto/nds/CotoND200305.pdf>
- [16] Cristancho C., Orlando (2004). Segmentación de imágenes digitales mediante morfología matemática. Consultado en Septiembre de 2009 en www.geocities.com/ocristanhoc/def_problema_occ.pdf
- [17] Delgado Palomino, Jorge Antonio (2002). Redes neuronales. Consultado en Abril de 2010 en <http://www.monografias.com/trabajos38/redes-neuronales/redes-neuronales.shtml>
- [18] Deselaers, T.; Ney, H.; Seidl, T.; Keysers, D. (2003). Features for Image Retrieval. Consultado en Junio 2010 en <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.126.241&rep=rep1&type=pdf>
- [19] Eakins, John; Graham, Margaret. Content-based Image Retrieval. Consultado en Febrero 2010 en http://www.jisc.ac.uk/uploaded_documents/jtap-039.doc
- [20] Florea, Rogozan, Cornea, Bensrhair, Darmoni (2006). MedIC/CISMeF at ImageCLEF 2006: Image Annotation and Retrieval Tasks. Proceedings of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.

- [21] Forsyth, D.; Ponce, J. (2003). Computer visión: A modern approach. Consultado en Abril de 2010.
- [22] G. Duarte, Oscar (2000). Sistemas de Lógica Difusa. Consultado en Septiembre de 2009 en <http://atenea.unicauca.edu.co/~wpantoja/web/inteligenciaartificial/fuzzy.pdf>
- [23] Gabor, D. (1946). Theory of Communication. Journal of Institute for Electrical Engineering. Consultado en Enero 2010 en <http://granularsynthesis.com/pdf/gabor.pdf>
- [24] García, Gustavo Iván (2007). Revisión de técnicas de agrupamiento de minería de datos espaciales en un SIG. Consultado en Noviembre de 2009 en <http://www.monografias.com/trabajos27/datamining/datamining.shtml#arquit>
- [25] González, Rafael C.; Woods, Richard E. (2008). Digital Image Processing. Consultado en Septiembre de 2009 en <http://books.google.com/books?id=8uGOnjRGEzoC&printsec=frontcover&dq=%22digital+image+processing%22+gonzalez&lr=&ei=ndyVSrGSDIPCmQKNkdCH#v=onepage&q&f=false>
- [26] González Díaz, Ernesto (2004). Técnicas de minería de datos. Consultado en Enero de 2010 en <http://www.monografias.com/trabajos55/mineria-de-datos/mineria-de-datos.shtml>
- [27] Grigorescu, Simona; Petkov, Nicolai; Kruizinga, Meter (2002). Comparison of texture features based on Gabor Filters. Consultado en Enero 2010 en <http://www.cs.rug.nl/users/petkov/publications/2002ieee-tip.pdf>
- [28] Gutiérrez, José Manuel (2003). Sistemas expertos basados en reglas. Consultado en Septiembre 2009 en <http://personales.unican.es/gutierjm/cursos/expertos/Reglas.pdf>
- [29] Güld MO, Deserno TM. (2007). Baseline Results for the CLEF 2007: Medical automatic annotation task. Working Notes of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.
- [30] Güld MO, Thies C, Fisher B, Lehmann TM. Baseline (2006). Combining global features within a nearest neighbour classifier for content-based retrieval of medical images. Proceedings of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.

- [31] Hersh W, Kalpathy-Cramer J, Jensen J. (2006). Medical image retrieval and automated annotation: OHSU at ImageCLEF 2006. Proceedings of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.
- [32] Hersh W, Kalpathy-Cramer J. (2007). Medical image retrieval and automated annotation: OHSU at ImageCLEF 2007. Working Notes of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.
- [33] John H., Adreae (1999). Associative Learning for a Robot Intelligence (Artificial Intelligence). Consultado en Septiembre de 2009.
- [34] K. Laws (1980). Textured Image Segmentation, Ph.D. Dissertation, University of Southern California. Consultado en Diciembre de 2009
- [35] Lana Serrano, S.; Villena Román, J.; González Cristóbal, J.C.; Goñi Menoyo, J.M. (2007). MIRACLE at ImageCLEF 2007: Machina Learning Experiments on Medical Image Annotation. Working Notes of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.
- [36] Lanzarini, Laura; Corbalán, Leonardo; Osella Massa, Germán; De Giusti, Armando (2005). Redes neuronales, aplicaciones en reconocimiento de patrones y diseño de agentes reflejos. Consultado el 20 de Marzo de 2009 en <https://dc.exa.unrc.edu.ar/wicc/papers/InteligenciaArtificial/36.pdf>
- [37] Lehman, Thomas M.; Schubert, Henning; Keysers, Daniel; Kohnen, Michael; Wein, Berthold B (2003). The IRMA code for unique classification of medical images. Consultado en Septiembre de 2009 en <http://libra.imib.rwth-aachen.de/irma/ps-pdf/Proceedings-Lehmann.pdf>.
- [38] Lehmann, Thomas M. (2004). IRMA: Image Retrieval in Medical Applications. Consultado en Octubre de 2009 en http://irma-project.org/index_en.php
- [39] Liu J, Hu Y, Li M, Ma WY (2006). Medical image annotation and retrieval using visual features. Proceedings of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.
- [40] Mallat, S.G. (1989). "Multifrequency Channel Decompositions of Images and Wavelet Models" IEEE Transactions on Acoustics, Speech, and Signal Processing, December 1989. Consultado en Abril de 2010.

- [41] McCulloch, Warren; Pitts, Walter (1943). A Logical Calculus of Ideas Immanent in Nervous Activity. Consultado en Noviembre de 2009.
- [42] Montes Castro, Jesús (2004). Sistemas Expertos. Consultado en Septiembre de 2009 en <http://www.monografias.com/trabajos16/sistemas-expertos/sistemas-expertos.shtml>
- [43] Newell, Allen y Simon, Herbert A. (1972). Human problem solving. Consultado en Septiembre de 2009 en http://www.sci.brooklyn.cuny.edu/~kopec/cis718/fall_2005/2/Rafique_2_humanthinking.doc.
- [44] P.C. Mahalanobis (1936). On the generalised distance in statistics, Proceedings of the National Institute of Science of India 12. Consultado en Febrero de 2010
- [45] Petkova D., Ballesteros L. (2005). Categorizing and annotation medical images by retrieving terms relevant to visual features. Consultado en Septiembre de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.9398&rep=rep1&type=pdf>.
- [46] P. Pentland (1984). Fractal-based description of natural scenes, IEEE Transactions on Pattern Analysis and Machine Intelligence. Consultado en Enero de 2010.
- [47] Pratt, W. K. (2001). Digital Image Processing, 3rd ed., Wiley Interscience, NY. Consultado en Octubre de 2009.
- [48] Priore, Paolo; De la Fuente, David; Pino Raúl y Puente, Javier (2002). Utilización de las Redes Neuronales en la toma de decisiones. Aplicación en un problema de secuenciación. Consultado en Octubre de 2009 en https://www.icaei.es/contenidos/publicaciones/anales_get.php?id=721
- [49] Qiu B, Xu C, Tian Q. (2006). Two-stage SVM for Medical Image Annotation. Proceedings of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.
- [50] Qiu B, Xiong W, Tian Q y Xu CS (2005). Report for annotation task in ImageCLEFmed 2005. Consultado en Octubre de 2009.
- [51] Rahman MM, Desai BC, Bhattacharya P. (2005). Supervised machine learning based medical image annotation and retrieval. Consultado en Octubre de 2009 en <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.106.9405&rep=rep1&type=pdf>.

- [52] Ríos-Díaz, José; Martínez-Payá, Jacinto Javier; del Baño Aledo, M^a Elena (2009). Textural Analysis by Means of a Grey Level Co-Occurrence Matrix Method on Patellar Tendon Ultrasonography is Useful for the Detection of Histological Changes after Whole-Body Vibration Training. Consultado en Abril de 2010 en <http://redalyc.uaemex.mx/pdf/1630/163013077004.pdf>
- [53] Rusell S, Norvig P. Inteligencia artificial. Un enfoque moderno. Ed. Pearson Prentice Hall. 2^a Edición. Consultado en Octubre de 2009.
- [54] Salazar Serrudo, Carla (2003). Una introducción a los Sistemas Expertos. Texto de apoyo para la asignatura Sistemas Expertos. Consultado en Septiembre de 2009 en <http://www.fcyt.umss.edu.bo/docentes/269/publicacion/IntSE.pdf>.
- [55] Setia, L.; Teynor, A.; Halawani, A.; Burkhardt, H. (2006). University of Freiburg at ImageCLEF06 - Radiograph annotation using local relational features. Proceedings of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.
- [56] Springmann M.; Schuldt, H. (2007). Speeding up IDM without degradation of retrieval quality. Working Notes of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.
- [57] Tagare, H.D; Jaffe, C.C y Duncan, J. (1997). Medical image databases - A content-based retrieval approach. Journal of the American Medical Informatics Association JAIMA. Consultado en Abril de 2010 en <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=61234>.
- [58] Tamura, H.; Mori, S.; Yamawaki, T. (1978). Textural features corresponding to visual perception. IEEE Transactions on Systems, Man and Cybernetics. Consultado en Abril de 2010.
- [59] Tomás, T.; Orabona, F.; Caputo, F (2007). CLEF2007 Image annotation task: an SVM-based cue integration approach. Working Notes of the Cross Language Evaluation Forum (CLEF). Consultado en Octubre de 2009.
- [60] Apuntes de la Universidad de Cantabria sobre Sistemas Expertos basados en probabilidad. Consultado en Octubre de 2009 en <http://www.docstoc.com/docs/3251073/SISTEMAS-EXPERTOS-BASADOS-EN-PROBABILIDAD-En-las-aplicaciones-la-incertidumbre>
- [61] <http://zeus.lci.ulsa.mx/divulgacion/Material/pdf/Reconocimiento%20de%20Patrones.pdf>. Consultado en Abril de 2010.

- [62] <http://www.clef-campaign.org/> (Consultado en repetidas ocasiones durante la realización del proyecto).
- [63] <http://www.daedalus.es/i-d-i/clef-cross-language-evaluation-forum/> Consultado en Diciembre de 2009.
- [64] <http://medical.nema.org>. Consultado en Noviembre de 2009.
- [65] http://www.tsc.uc3m.es/imagen/Curso_ProcesadoMorfologico/Contenido/Operaciones/OperacionesMorfologicas.html Consultado en Febrero de 2010.
- [66] <http://www.mathworks.es/products/matlab> Consultado en Junio de 2010.
- [67] Transformadas de imágenes:
http://www.diac.upm.es/acceso_profesores/asignaturas/tdi/tdi/transformadas/transformadas.htm Consultado en Enero de 2010.
- [68] <http://dicom.nema.org> Consultado en Diciembre de 2009.
- [69] FIRE: <http://www-i6.informatik.rwth-aachen.de/~deselaers/fire/> Consultado en Junio de 2010.